# DEMONSTRATION REPORT

## Data Modeling, Feature Extraction, and Classification of Magnetic and EMI Data, ESTCP Discrimination Study, Camp Sibert, AL

## ESTCP Project MM-0504

SEPTEMBER 2008

Dr. Stephen Billings
Sky Research, Inc.
.

Environmental Security Technology Certification Program

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | 9/5/2008 | Final, 2007-2008 |

**4. TITLE AND SUBTITLE**
Data Modeling, Feature Extraction, and Classification of Magnetic and EMI Data,
ESTCP Discrimination Study, Camp Sibert, AL
Project 200504: Practical Discrimination Strategies for Application to Live Sites

**5. FUNDING NUMBERS**
Contract # W912HQ-05-C-0018

**6. AUTHOR(S)**
Dr. Stephen Billings

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Sky Research, Inc.
445 Dead Indian Memorial Rd.
Ashland, OR 97520

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
ESTCP Program Office
901 North Stuart Street, Ste 303
Arlington, VA 22203-1821

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
Unclassified/Unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT (Maximum 200 words)**

An ESTCP-sponsored demonstration was conducted at the Former Camp Sibert, Alabama, where the objective was to discriminate potentially hazardous 4.2" mortars from non-hazardous shrapnel, range and cultural debris. Nine different discrimination techniques were tested and used data from the MTADS magnetometer array, a Geonics EM61 cart, the MTADS EM61 array and the Geonics EM63. Discrimination was achieved using either rule-based or statistical classification of feature vectors extracted from dipole or polarization tensor model fits to detected anomalies. Each of the EM methods utilized a size and a time-decay based feature vector. All nine different methods were successful at discriminating the 4.2" mortars from non-hazardous items. The EM based methods were more effective than the magnetometer. The higher SNR, denser coverage and more precise positioning of the MTADS EM61 array data resulted in fewer false-positives than the EM61 cart. When depth constraints from the magnetometer data were used to constrain the EM fits through cooperative inversion, discrimination performance was improved. The most effective discrimination technique used the cooperatively inverted fits to the EM63 cued interrogation data. All UXO were recovered before a single false-positive was excavated.

**14. SUBJECT TERMS**
UXO, discrimination, EM, magnetometry

**15. NUMBER OF PAGES**
111

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | Unlimited |

## EXECUTIVE SUMMARY

The demonstration described in this report was conducted at the Former Camp Sibert, Alabama, under project ESTCP MM-0504 "Practical Discrimination Strategies for Application to Live Sites." It was performed under the umbrella of the ESTCP Discrimination Study Pilot Program. The MM-0504 project is attempting to demonstrate the application of feature extraction and statistical classification to the problem of UXO discrimination. At the Camp Sibert site the objective was to discriminate potentially hazardous 4.2" mortars from non-hazardous shrapnel, range and cultural debris. In this report, we describe the performance of nine different discrimination techniques that utilized data from the MTADS magnetometer array, a Geonics EM61 cart, the MTADS EM61 array and the Geonics EM63.

Each of the discrimination techniques utilized features extracted from a phenomenological model that was fit to the observed data around each anomaly. For magnetics, the model was a static dipole, while for EM a polarization tensor model was used. Location constraints from the magnetometer were used to constrain the MTADS EM61 and EM63 models in a cooperative inversion procedure. From the extracted feature vectors the following nine different prioritized dig-lists were created: (i) Magnetics ranked by moment; (ii) Magnetics ranked by remanence; (iii) EM61 ranked by maximum polarizability; (iv) EM61 statistical classification; (v) MTADS EM61 ranked by maximum polarizability; (vi) MTADS EM61 statistical classification; (vii) MTADS EM61 cooperative inversion and statistical classification; (viii) EM63 statistical classification; and (ix) EM63 cooperative inversion and statistical classification. For statistical classification, a Probabilistic Neural Net was trained on a size and a time-decay feature extracted from anomalies in the Geophysical Prove Out and a partial release of ground truth. Stop digging points were selected by professional judgment. All model fits and statistical classification were performed using the UXOLab software that was jointly developed by UBC-GIF and Sky Research, principally through funding from the USACE-ERDC.

One-hundred and nineteen seeded 4.2" mortars were left in the blind-test dataset. Each of the nine different methods was successful at discriminating the 4.2" mortars from non-hazardous items. For magnetics, the ranking by moment was better than that of remanence due to relatively large remanent magnetizations of several of the seeded items. At the selected operating point 76% of non-hazardous items were left in the ground and all 4.2" mortars were recovered. For the EM methods, the rankings by statistical classification were better than those based on size alone, with further improvements possible when using the cooperatively inverted feature vectors. Statistical classification of the contractor-collected EM61 data resulted in recovery of all 4.2" mortars with 59% of non-hazardous items left in the ground. For the MTADS EM61 and MTADS EM61 cooperative processes, all mortars were recovered with 52% and 72%, respectively, of non-hazardous items unexcavated. The MTADS EM61 results were degraded by the large number of "can't analyze" anomalies that had to be excavated as suspected UXO. Many of these were geological artifacts that were caused by cart-bounce as the MTADS EM61 array traversed perpendicular furrows in a recently plowed field. Retrospective analysis revealed that many of these "can't analyze" anomalies could have been eliminated using either a Figure of

Merit or a geological pre-screener. The pre-screener was based on a statistical classification of feature vectors related to the energy in the first time-channel in the east-west versus north-south transects of the MTADS EM61 array. Over 55% of the geological anomalies could be rejected because they had significantly more energy north-south (perpendicular to the furrows), than east-west (parallel to the furrows).

Excluding the "can't analyze" category revealed the excellent intrinsic discrimination ability of the EM methods, particularly the MTADS EM61 cooperative. All of the one-hundred and eighteen 4.2" mortars were recovered with just 3 false-positives, compared to 75, 25 and 8 false-positives for the magnetometer, EM61 and MTADS EM61 respectively. Stop-digging points were very conservative with 103, 115, 59 and 70 false-positives for the magnetometer, EM61, MTADS EM61 and MTADS EM61 cooperative, respectively.

The EM63 was deployed in a cued-interrogation mode over a subset of anomalies: the blind-test data comprised 150 items which contained 34 UXO. At the selected operating point, the EM63 statistical classification required 38 false-positives but one 4.2" mortar was missed. The failure was caused by a corrupt line of data that resulted in a small, shallow solution with a lower misfit than an alternative second solution closer to the true depth of burial. When the depth from the magnetometer was used as a constraint in the cooperative inversion process, the false-negative did not occur. In fact, the EM63 cooperative produced a "perfect" ROC curve, with all 34 UXO recovered with 0 false-alarms. At the selected operating point 21 false-positives were required, with 16 of these in the "can't analyze" category.

The depth predictions of the magnetometer data are very accurate with over 85% within 10 cm and 96% within 20 cm of the true depths. The EM61 and MTADS EM61 display a much larger scatter in actual and predicted depths with 74 and 76% respectively within 20 cm. For the EM61, the estimated depths of the deeper 4.2" mortars are particularly poor and contribute to the relatively wide range of estimated size observed for that class. Both the EM61 and MTADS EM61 have a tendency to predict deeper depths for small, shallow items. This characteristic was also exhibited by the EM61 and EM63 data at FLBGR during the last demonstration conducted under this ESTCP project.

Cooperative inversion considerably improves the accuracy of the estimated depths: 75% of MTADS EM61 depths are within 10 cm compared to 53% when inverted without magnetometer depth constraints. The performance gain is from 75% to 87% within 10 cm when the EM63 data are cooperatively inverted. The better-constrained depths of the cooperatively inverted EM models result in less scatter in the polarizabilities and an improved discrimination ability.

# TABLE OF CONTENTS

# LIST OF TABLES

## LIST OF FIGURES

## ACRONYMS

| | |
|---|---|
| APG | Aberdeen Proving Ground |
| ASR | Archive Search Report |
| BOR | Body of Revolution |
| COTS | Commercial off-the-shelf |
| Cs | Cesium |
| EM | Electromagnetic |
| EMI | Electromagnetic induction |
| ERDC | Engineer Research & Development Center |
| ESTCP | Environmental Security Technology Certification Program |
| FLBGR | Former Lowry Bombing and Gunnery Range |
| FOM | Figure of Merit |
| FP | False Positive |
| FUDS | Formerly Used Defense Site |
| GPO | Geophysical Prove-Out |
| HE | High Explosive |
| IDA | Institute for Defense Analysis |
| IMU | Inertial Measurement Unit |
| MEC | Munitions and explosives of concern |
| MR | Munitions Response |
| MTADS | Multi-Sensor Towed Array Detection System |
| NA | Not Applicable |
| NRL | Naval Research Laboratory |
| nT | NanoTesla |
| Pd | Probability of Detection |
| PDisc | Probability of Discrimination |
| PI | Principal Investigator |
| Pfa | Probability of False Alarm |
| PNN | Probabilistic Neural Network |
| QA | Quality Assurance |
| QC | Quality Control |
| ROC | Receiver operating characteristic |
| RR | Rocket Range |
| RTC | Replacement Training Center |
| RTS | Robotic Total Station |
| SE1 | South-East-1 |
| SE2 | South-East-2 |
| SERDP | Strategic Environmental Research and Development Program |
| SNR | Signal to noise ratio |
| SVM | Support vector machine |
| SW | South-West |
| TEM | Time-domain electromagnetic |
| TEU | Technical Escort Unit |
| UBC-GIF | University of British Columbia – Geophysical Inversion Facility |
| YPG | Yuma Proving Ground |

| | |
|---|---|
| USACE | United States Army Corps of Engineers |
| UTC | Unit Training Center |
| UXO | Unexploded ordnance |
| WP | White Phosphorus |

## ACKNOWLEDGEMENTS

The work described in this demonstration report was performed jointly by Sky Research Inc. and the University of British Columbia-Geophysical Inversion Facility. All work was conduced at either the Sky Research Vancouver office, or at the UBC-GIF office which is also in Vancouver. Dr. Stephen Billings served as Principal Investigator.

Funding for this project was provided by the Environmental Security Technology Certification Program Office. This project offered the opportunity to demonstrate this technology and evaluate its potential use to support the Department of Defense's efforts to discriminate hazardous ordnance from non-hazardous shrapnel, ordnance related scrap and cultural debris.

We wish to express our sincere appreciation to Dr. Jeffrey Marqusee, Dr. Anne Andrews, and Ms. Katherine Kaye of the ESTCP Program Office for providing support and funding for this project, and to Dr Herbert Nelson of the Naval Research Laboratory for facilitating access to the technical information required to complete this demonstration. We also would like to acknowledge the support and direction provided by Mr. Don Yule of the USACE-ERDC, the COR for this project.

We also would like to acknowledge the long term support and funding provided by the USACE-ERDC and Dr John Cullinane in particular. The work reported here is a demonstration of techniques largely funded through the USACE-ERDC.

# 1 INTRODUCTION

## 1.1 Background

The FY06 Defense Appropriation contains funding for the "Development of Advanced, Sophisticated, Discrimination Technologies for UXO Cleanup" in the Environmental Security Technology Certification Program (ESTCP). In 2003, the Defense Science Board observed: "The … problem is that instruments that can detect the buried unexploded ordnance (UXO) also detect numerous scrap metal objects and other artifacts, which leads to an enormous amount of expensive digging. Typically 100 holes may be dug before a real UXO is unearthed! The Task Force assessment is that much of this wasteful digging can be eliminated by the use of more advanced technology instruments that exploit modern digital processing and advanced multi-mode sensors to achieve an improved level of discrimination of scrap from UXO."

Significant progress has been made in discrimination technology. To date, testing of these approaches has been primarily limited to test sites, with only limited application at live sites. Acceptance of discrimination technologies requires demonstration of system capabilities at real UXO sites under real world conditions. Any attempt to declare detected anomalies to be harmless and requiring no further investigation will require demonstration to regulators of not only the effectiveness of individual technologies, but an entire decision making process. This discrimination study is the first phase in what is expected to be a continuing effort that will span several years.

## 1.2 Objectives of the Demonstration

### 1.2.1 Objectives of the ESTCP UXO Discrimination Study

As outlined in the Environmental Security Technology Certification Program Unexploded Ordnance Discrimination Study Demonstration Plan (ESTCP, 2006), the objectives of the study are twofold. First, the study was designed to test and validate UXO detection and discrimination capabilities of currently available and emerging technologies on real sites under operational conditions. Second, the ESTCP Program Office and their demonstrators are investigating, in cooperation with regulators and program managers, how UXO discrimination technologies can be implemented in cleanup operations.

Within each of these two overarching objectives, there are several sub-objectives.

### 1.2.2 Technical objectives of the Discrimination Study

The study was designed to test and evaluate the capabilities of various UXO discrimination processes which each consist of selected sensor hardware, a survey mode, and a software-based processing step. These advanced methods can then be compared to existing practices and be used to validate the pilot technologies for the following:

- Detection of UXO
- Identification of features that can help distinguish scrap and other clutter from UXO

- Reduction of false alarms (items that could be safely left in the ground that are incorrectly classified as UXO) while maintaining Probabilities of Detection ($P_d$s) acceptable to all
- Quantify the cost and time impact of advanced methods on the overall cleanup process as compared to existing practices

Additionally, the study aims to understand the applicability and limitations of the selected technologies in the context of project objectives, site characteristics, and suspected ordnance contamination. Sources of uncertainty in the discrimination process were identified and their impact quantified to support decision making. This included issues such as impact of data quality due to how the data are collected. The process for making the dig-no dig decision process was explored. Potential quality assurance/quality control (QA/QC) processes for discrimination also were explored. Finally, high-quality, well documented data were collected to support the next generation of signal processing research.

## 1.3 Regulatory Drivers and Stakeholder Issues

ESTCP has assembled an Advisory Group to address the regulatory, programmatic, and stakeholder acceptance issues associated with the implementation of discrimination in the Munitions Response (MR) process.

### 1.3.1 Objective Of The Advisory Group

The advisory group will focus on exploring UXO discrimination processes that will be useful to regulators and site managers in making decisions by determining:

- What information is required to support a discrimination decision?
- What data are needed to support decisions, particularly with regard to decisions not to dig all detected anomalies?
- Necessary end-products to support discrimination decisions.
- What are the site-specific factors that impact this process?
  - How best can the information be presented?
- What must be demonstrated for the regulatory community to consider not digging every anomaly as a viable alternative?
  - Methodology
  - Transparency
  - QA/QC requirements
  - Validation
- For implementation beyond the pilot project, how should proposals to implement discrimination be evaluated?

In support of the objective stated above, the advisory group will provide input and guidance to the Program Office on the following topics:

- Pilot project objectives and flow-down to performance metrics
- Flow down of program objectives to data quality objectives
- Demonstration / data collection plans
- QA/QC requirements and documentation
- Interpretation, analysis, and validation
- Process flow for discrimination-based removal actions
- What does it all mean?

### 1.3.2 Specific Objectives of the Demonstration

The objectives of this demonstration within the larger discrimination study objectives were to perform data modeling, classification and discrimination, using magnetometer and electromagnetometer (EM) data collected by the various demonstrators participating in the study. Specifically, we performed the following:

1) Data modeling:
   a) Dipole fitting of the magnetometer data;
   b) Fitting of 2- and 3-dipole beta models to the Multi-Sensor Towed Array Detection System (MTADS) EM61 data and selection of optimal models;
   c) Fitting of 2- and 3-dipole Pasion-Oldenburg models to the EM63 cued-interrogation data and selection of optimal models; and
   d) Cooperative inversion of the EM61 and EM63 data using the dipole fits from the magnetometer data to constrain the object's location and depth;
2) Classification and discrimination:
   a) Magnetics size-based: Calculation of the magnetic remanence metric and the production of a dig-sheets ranked according to moment and one according to remanence;
   b) MTADS EM61 size-based: Production of a dig-sheet ranked according to size (using the sum of the beta parameters for time-channel 1 from the MTADS EM61 data).
   c) EM61 size-based (Contractor): Production of a dig-sheet ranked according to size (using the sum of the beta parameters for time-channel 1 from the Contractor EM61 data).
   d) MTADS EM61 statistical: Statistical classification of features derived from the MTADS EM61 data and the production of a ranked dig-sheet;
   e) EM61 statistical (Contractor): Statistical classification of features derived from the Contractor EM61 data and the production of a ranked dig-sheet;
   f) EM63 statistical: Same as b) but with the EM63;
   g) MTADS EM61 and magnetics statistical: As per b) but with MTADS EM61 fits constrained by the magnetics data and with the addition of the features from the magnetometer data (remanence, moment etc); and
   h) EM63 and magnetics statistical: As per d) but with the EM63.

We had planned to produce 8-ranked dig-sheets but in the end produced 9 (we submitted two for the magnetics).

The first demonstration of the methodology defined in this research project was conducted at the Former Lowry Bombing and Gunnery Range (FLBGR) in Colorado during the 2006 field season. The focus of the FLBGR demonstration was on verification of the single inversion process used to extract physics-based parameters from magnetic and electromagnetic induction (EMI) anomalies, as well as the statistical classification algorithms used to make discrimination decisions from those parameters. This demonstration provided another test of that methodology as well as that of the cooperative inversion process.

# 2 TECHNOLOGY DESCRIPTION

## 2.1 Technology Development and Application

Magnetic and electromagnetic methods represent the main sensor types used for detection of UXO. Over the past 10 years, significant research effort has been focused on developing methods to discriminate between hazardous UXO and non-hazardous scrap metal, shrapnel and geology (e.g. Hart *et al*., 2001; Collins *et al*., 2001; Pasion & Oldenburg, 2001; Zhang *et al*., 2003a, 2003b; Billings, 2004). The most promising discrimination methods typically proceed by first recovering a set of parameters that specify a physics based model of the object being interrogated. For example, in time-domain electromagnetic (TEM) data, the parameters comprise the object's location and the polarization tensor (typically two or three collocated orthogonal dipoles along with their orientation and some parameterization of the time-decay curve). For magnetics, the physics based model is generally a static magnetic dipole. Once the parameters are recovered by inversion, a subset of the parameters is used as feature vectors to guide either a statistical or rule-based classifier.

Magnetic and EM phenomenologies have different strengths and weaknesses. Magnetic data are simpler to collect, are mostly immune to sensor orientation and are better able to detect deeper targets. EM data are sensitive to non-ferrous metals, are better at detecting smaller items and are able to be used in areas with magnetic geology. Therefore, there are significant advantages in collecting both types of data including increased detection, stabilization of the EM inversions by cooperative inversion of the magnetics (Pasion *et al*., 2003) and extra dimensionality in the feature space that may improve classification performance (e.g. Zhang *et al*., 2003a). However, these advantages need to be weighed against the extra costs of collecting both data types.

There are three key elements that impact the success of the UXO discrimination process described in the previous paragraphs:

1) Creation of a map of the geophysical sensor data: This includes all actions required to form an estimate of the geophysical quantity in question (magnetic field in [nano-Tesla] nT, amplitude of EMI response at a given time-channel, etc.) at each of the visited locations. The estimated quantity is dependent on the following:
    a. Hardware, including the sensor type, deployment platform, position and orientation system and the data acquisition system used to record and time-stamp the different sensors;
    b. Survey parameters such as line spacing, sampling rate, calibration procedures etc.;
    c. Data processing such as merging of position/orientation information with sensor data, noise and background filtering applied;
    d. The background environment including geology, vegetation, topography, cultural features, etc.; and
    e. Depth and distribution of ordnance and clutter.

2) Anomaly selection and feature extraction: This includes the detection of anomalous regions and the subsequent extraction of a dipole (magnetics) or polarization tensor (TEM) model for each anomaly. Where magnetic and EMI data have both been collected, the magnetic data were used as constraints for the EMI model via a cooperative inversion process.

3) Classification of anomalies: The final objective of the demonstration was the production of a dig sheet with a ranked list of anomalies. This was achieved via statistical classification which required training data to determine the attributes of the UXO and non-UXO classes.

The focus of this demonstration was on the further testing and validation of the methodologies for 2) and 3) above that have been developed in UXOLab jointly by Sky Research and the University of British Columbia-Geophysical Inversion Facility (UBC-GIF). The success of the discrimination process was critically dependent on the attributes of the data used for the feature extraction and subsequent classification (*vis-a-vis*, everything pertaining to the first element described above), in particular, the signal-to-noise ratio (SNR), location accuracy, sampling density and information content of the data (the more time channels or vector components, the more information that will be available to constrain the fits). Thus, while our intent was to test the algorithms developed in UXOLab, this test could not be conducted in isolation of the attributes of the geophysical sensor data. The inclusion of both the contractor and MTADS EM61 data, and the cued-interrogation data from the EM63 allowed the impact of data quality to be assessed.

Data to be used for the discrimination and classification were provided by the other Discrimination Study performers.

We now describe each of the three key elements of the technology as identified above.

### 2.1.1 Creation of a Map of Geophysical Sensor Data

Each of the demonstrators provided filtered, located geophysical data. No additional pre-processing was performed on any of the datasets.

### 2.1.2 Anomaly Selection and Feature Extraction

At this point in the process flow, there was a map of each of the geophysical quantities measured during the survey. The next step in the process was detection of anomalous regions followed by the extraction of features for each of the detected items.

*Feature Extraction: Time-domain Sensor*

In the EMI method, a time varying field illuminates a buried, conductive target. Currents induced in the target then produce a secondary field that is measured at the surface. EM data inversion involves using the secondary field generated by the target for recovery of the position, orientation, and parameters related to the target's material properties and shape. In the UXO community, the inverse problem is simplified by assuming that the secondary field can be accurately approximated as a dipole.

In general, TEM sensors use a step off field to illuminate a buried target. The currents induced in the buried target decay with time, generating a decaying secondary field that is measured at the surface. The time-varying secondary magnetic field **B**(*t*) at a location **r** from the dipole **m**(*t*) is:

$$\mathbf{B}(t) = \frac{\mu_o}{4\pi r^3}\mathbf{m}(t)\cdot\left(3\hat{\mathbf{r}}\hat{\mathbf{r}} - \bar{\bar{\mathbf{I}}}\right) \tag{1}$$

where $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$ is the unit-vector pointing from the dipole to the observation point, $I$ is the 3 x 3 identity matrix, $\mu_o = 4\ \pi$ x $10^{-7}$ H/m is the permittivity of free space and $r = |\mathbf{r}|$ is the distance between the center of the object and the observation point.

The dipole induced by the interaction of the primary field $\mathbf{B}_o$ and the buried target is given by:

$$\mathbf{m}(t) = \frac{1}{\mu_o}\bar{\bar{\mathbf{M}}}(t)\cdot\mathbf{B}_o \tag{2}$$

where **M**(*t*) is the target's polarization tensor. The polarization tensor governs the decay characteristics of the buried target and is a function of the shape, size, and material properties of the target. The polarization tensor is written as:

$$\bar{\bar{\mathbf{M}}}(t) = \begin{bmatrix} L_1(t) & 0 & 0 \\ 0 & L_2(t) & 0 \\ 0 & 0 & L_3(t) \end{bmatrix} \tag{3}$$

where we use the convention that $L_1(t_1) \geq L_2(t_1) \geq L_3(t_1)$, so that polarization tensor parameters are organized from largest to smallest. The polarization tensor components are parameterized such that the target response can be written as a function of a model vector containing components that are a function of target characteristics. Particular parameterizations differ depending on the instrument (number of time channels, time range measured etc) and the group implementing the work. Bell *et al*. (2001) solves for the components of the polarization tensor at each time channel, and this is the procedure used for the four channel Geonics EM61 MKII. For the EM63 we used the Pasion-Oldenburg formulation (Pasion and Oldenburg, 2001):

$$L_i(t) = k_i(t + \alpha_i)^{-\beta_i}\exp(-t/\gamma_i) \tag{4}$$

for i={1,2,3},with the convention that $k_1 \geq k_2 \geq k_3$. For a body-of-revolution (BOR), $L_2 = L_3$ for a rod-like object (Pasion and Oldenburg, 2001) and $L_1 = L_2$ for a plate-like object.

Given a set of observations $\mathbf{d}^{\mathbf{obs}}$, we formulate the parameter estimation as an optimization problem through Bayes theorem:

$$p(\mathbf{m}\,|\,\mathbf{d}^{\mathbf{obs}}) = \frac{p(\mathbf{m})p(\mathbf{d}^{\mathbf{obs}}\,|\,\mathbf{m})}{p(\mathbf{d}^{\mathbf{obs}})} \tag{5}$$

where **m** is the vector of model parameters (location, orientation and polarization tensor parameters), $p(\mathbf{m})$ is the probability distribution representing prior information, $p(\mathbf{d}^{\mathbf{obs}})$ is the marginal probability density of the experimental data, and $p(\mathbf{d}^{\mathbf{obs}}|\mathbf{m})$ is the conditional probability density of the experimental data which describes the ability of the model to reproduce the experimental data. The *a-posteriori* conditional probability density $p(\mathbf{m}|\mathbf{d}^{\mathbf{obs}})$ is the probability

density we ascribe to **m** after collecting the data. The *a-posteriori* conditional probability density encapsulates all the information we have on the model parameters and the model that maximizes it is usually regarded as the solution to the inverse problem. We estimate a value of **m** that maximizes the log of the *a-posteriori* conditional probability density:

$$\mathbf{m}^* = \max_{\mathbf{m}} \left\{ \log\left( p\left( \mathbf{m} \mid \mathbf{d}^{\mathbf{obs}} \right) \right) \right\} \tag{6}$$

With a single data-set and no prior information on the model parameters (except maybe some bound constraints on the model parameters):

$$\text{minimize } \phi(\mathbf{m}) = \frac{1}{2} \left\| V_d^{-1/2} \left( \mathbf{d}^{\mathbf{obs}} - F(\mathbf{m}) \right) \right\|^2, \text{subject to } m_i^L \le m_i \le m_i^U . \tag{7}$$

where $F(\mathbf{m})$ is a vector comprising the forward modeled data at the sampled locations, $m_i^L$ and $m_i^U$ are the lower and upper bounds on parameter $i$ and $V_d$ is the co-variance matrix of the data. Efficient algorithms for the solution of this optimization problem have been implemented for various polarization tensor formulations within UXOLab (including two- and three independent polarization tensors).

In this section we outline our choices for these components when applying parameter estimation to the Camp Sibert data sets.

**Data Covariance Matrix** $V_d$

Our knowledge of the noise levels appropriate to the solution of the inverse problem is encapsulated in the data covariance matrix. We assume independently distributed Gaussian errors and use the following data covariance matrix:

$$\left[ V_d^{-1/2} \right]_{ij} = \begin{Bmatrix} 0 & \text{if } i \neq j \\ \dfrac{1}{\delta_i + \varepsilon_i} & \text{if } i = j \end{Bmatrix} \tag{8}$$

where $\delta_t$ is a percentage of the $i^{th}$ datum:

$$\delta_i = \%error \times \left[ d_{obs} \right]_i \tag{9}$$

and $\varepsilon_i$ is a base level error that is present in the $i^{th}$ datum in the absence of a target. In the previous demonstration we assumed 0 %error, so the data covariance matrix was constant in space and determined by the base level errors, which are a function of time. Our strategy for determining the base line error was to mask a section of data within a grid that has no targets, and calculate the data statistics for each time channel within that grid. A single value represented the baseline error for the entire grid. However, after further analysis we found that noise characteristics often changed significantly with survey event (different data collection periods). For this demonstration we made two significant variations to the covariance matrix estimation:

1. We included a non-zero %error term as our simulations have shown that inaccurate positioning contributes an error that is reasonably approximated by an appropriate %error;

2. For EMI sensor data we used a new background estimation procedure to approximate the base-error level as a function of spatial position and time.

Background Estimation

When determining both the mean and standard deviation of a noise distribution that is hidden within a combined response of an anomalous signal, we make two assumptions.

1) The distribution of noise must be symmetric about its mean; and
2) Signals of interest are positive and hence influence the upper side of the mean value of the noise distribution.

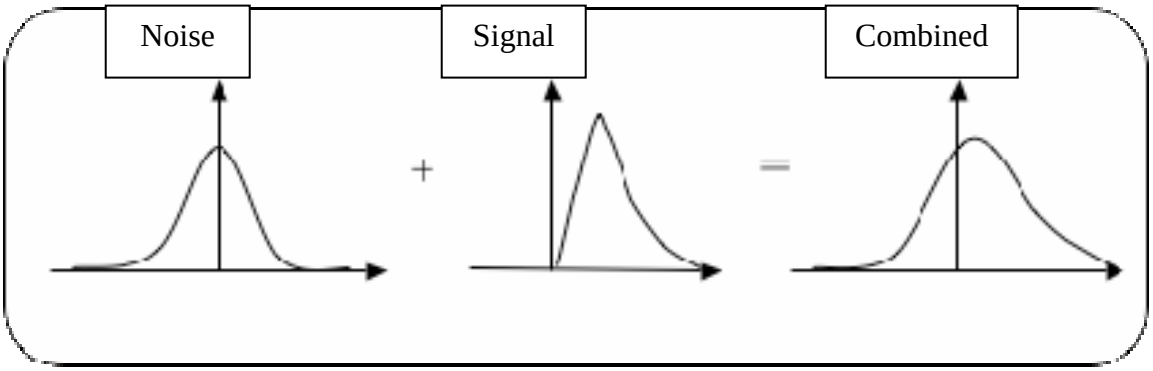



**Figure 1.** Graphic illustration showing the superposition of the background noise and anomalous signal distributions.

In practice the combined distribution of background noise and signal will be the superposition of a bell-shaped and log-normal distribution respectively (Figure 1). The combined distributions will be 'tangled' or mixed such that it will be impossible to separate, therefore any estimation of the background noise's mean or standard deviation will be biased towards the mean of the anomalous signal.

Using a sliding window that samples the entire data region, the following steps are performed at each sampling location, thus providing an estimate of the current windows mean and standard deviation.

1) Select a spatial region of data where the mean and standard deviation are expected to remain constant in value. The region should be large enough to sample both signal and noise.
2) Sort the selected data from largest to smallest.
3) Iterate the following steps.
   a) Determine the skewness value, $\gamma_i$, of the current subset of data. For the first iteration this is just the data selected from the region in step 1.
   b) Determine the sign of the skewness (where $\gamma_{thresh}$ is a positive user defined value)
      i) For positive skewness, $\gamma_i > \gamma_{thresh}$
      ii) For negative skewness, $\gamma_i < -\gamma_{thresh}$
      iii) For zero skewness, $-\gamma_{thresh} \le \gamma_i \le \gamma_{thresh}$

c) If i) is true, then remove the maximum or if ii) is true remove the minimum and return to a). If iii) is true then proceed to step 4).

4) Compute both the mean and standard deviation from the remaining subset of data. This is the final estimate of both quantities.

Simulations and tests on real-data indicate that a default value of $\gamma_{thresh} = 0.25$ produces good results.

### Forming the Data Vector $\mathbf{d_{obs}}$

The inversion procedure assumes that we are dealing with a single target in free space. Sensor drift, background geology, and nearby targets are non-random errors in the data that bias the estimated polarization parameters. By appropriately detrending the data and masking the individual anomalies we can minimize these effects.

Defining the Data to be Inverted 1: Spatial coverage

Once data anomalies are identified, a mask is defined that represents the spatial limits of the data to be inverted. Unlike magnetics data, an unconstrained EMI inversion is very sensitive to adjacent anomalies and to the size of the mask used in areas without nearby anomalies. The masking procedure helps ensure that signal from adjacent anomalies does not affect the inversion results. In addition, from a practical standpoint, inverting the minimum number of observations reduces the computational time.

We employ an advanced masking procedure, which fits an ellipse to contours of the anomalous target. By using an ellipse we recover a relatively smooth-shaped mask that mimics the shape of the anomaly. The main challenge is to find contours that are both smooth and close to the noise level. Including our background estimates ensures that we choose appropriate starting contour values that are both above the baseline error and that encompass all of the anomalous data.

Defining the Data to be Inverted 2: Time Channels

For the FLBGR demonstration, we excluded any channels with a SNR of less than 10 dB from the inversion, while in this demonstration we included time-channels down to an SNR limit of 2 dB.

### Defining the Model $F[\mathbf{m}]$

Determining if a double-peaked anomaly should be inverted as a single target or a pair of targets

For this demonstration all anomalies were assumed to be due to a single target. If there was any uncertainty, the anomaly was placed in the "can't analyze" category.

Determining if a single target should be inverted for 2 or 3 unique polarizations

During the last demonstration, we performed both two and three-dipole fits and then compared misfits. If the data misfit of the three polarization inversion was less than 0.85 times the misfit of the two polarization inversion, then we selected the three polarization model. Implementation of this procedure was quite cumbersome as we had to invert and QC each model type separately, and then run a batch file outside of UXOLab to produce a final merged data file. In addition, the

features derived from the primary polarization provided the most useful discrimination information so that the model used generally had little impact on the final classification. For this demonstration, we found that the parameters of the primary polarization were highly diagnostic, and therefore fit all anomalies with a 3-dipole model.

Determining the parameterization of the polarization decay

There are a number of different techniques for parameterization of the temporal behavior of the polarization tensor. One common approach is to solve for the polarization value at each time channel (for example the AETC beta model which we refer to as the instantaneous amplitude polarization model). We apply this approach to EM61 data. This approach is less efficient for the 26 time-channels recorded by the EM63. For that sensor we parameterize the polarization decay. The parameterization is inspired by the different decay regimes observed in compact targets. At very early times, the decay of the voltage will follow a $t^{-1/2}$ decay, followed by a steeper power law decay ($t^{-3/2}$ for a sphere). At the late stage of the response decays exponentially. For this study we use the following parameterized version of the polarization decay for the Geonics EM63 data:

$$L(t) = kt^{-\beta} \exp\left(-\frac{t}{\gamma}\right) \tag{10}$$

**Optimization: Determining the minimum of $\phi$(m)**

The optimization routine we use for inversion is a local Newton-type method that minimizes the least squares objective/misfit. We address the problem of local minima and assess the level of ambiguity in resolving the depth of an item by choosing multiple starting models. We start each inversion by scanning the subsurface ($x$, $y$, $z$) up to a 1.2 m depth. At each position we solve for the non-diagonalized polarization tensor[1] for the first time channel (chosen for its superior signal-to-noise ratio). For each combination of a position and polarization tensor we compute a data misfit. The depth-misfit curve is defined by the best fit at a given depth (Figure 2, solid line). Starting models for the full inversion of multi-channel data are selected along the depth-misfit curve among the models with relative misfit below a given threshold, here 15% (circles). If the depth-misfit curve contains local minima these are also selected as starting models.

The iterative Newton-type inversion then proceeds with each starting model. A given search stops when the iteration reaches a set threshold (misfit tolerance or number of iterations). A final model is obtained for each of the starting models (black stars in Figure 2, note different misfit because computed on all time channels). In the example of Figure 2a there are final solutions with similar misfit spread over a 0.46 m depth range, which confirms the uncertainty in recovering depth. For comparison we show in Figure 2b the depth-misfit curve for a different target, where the minimum misfit is well defined as a function of depth, and therefore the depth is accurately recovered.

*Feature Extraction: Magnetics*

For magnetics, the physics-based model most commonly used is a dipole:

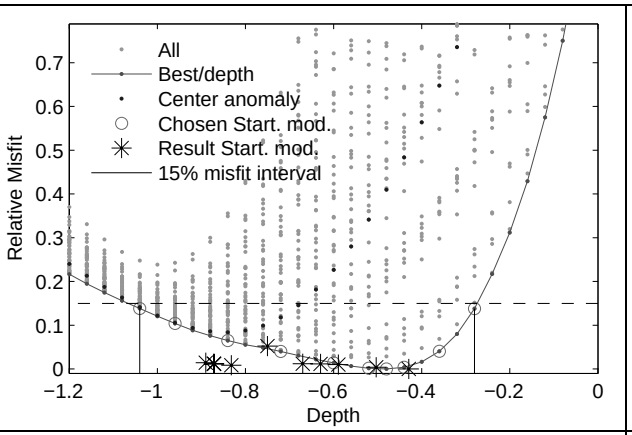
[1] When the polarization tensor is not explicitly diagonalized, the inverse problem is linear

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_O}{4\pi r^3}\mathbf{m}\cdot\left(3\hat{\mathbf{r}}\hat{\mathbf{r}} - \bar{\bar{\mathbf{I}}}\right) \tag{11}$$

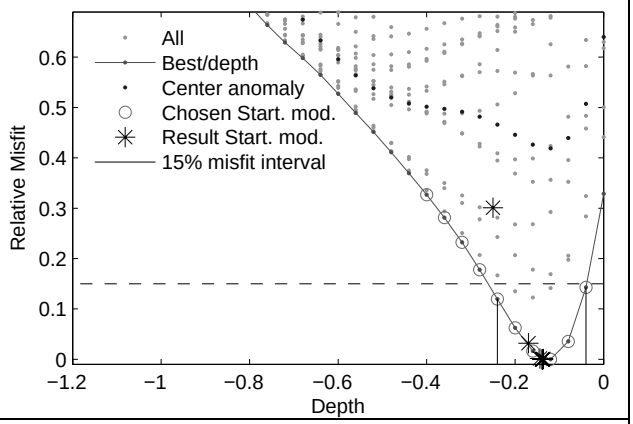
where the terms were defined earlier. As for the TEM case, a bound-constrained optimization problem is solved to extract feature vectors from each anomaly.

The magnetic remanence metric was calculated for each dipole moment (Billings 2004) by using an equivalent spheroid to represent the 4.2" mortar. The equivalent spheroid defines a dipole feasibility curve (the family of moments that an object of that length and diameter can produce by induced magnetization in the presence of the Earth's magnetic field). For each item, *i*, in the library (in this case there was only one item), the orientation is calculated that causes the minimum difference, $\mathbf{\Delta m_i}$, between the moment of the ordnance and the moment determined through the inversion, **m**. For each item in the library we then estimated the *minimum percentage of remanent magnetization* required to best match the observed dipole:

$$\mathbf{rem_i} = 100\frac{\|\mathbf{\Delta m_i}\|}{\|\mathbf{m}\|} \tag{12}$$




**a**. Depth-misfit relationship for the inversion of a 4.2" mortar. Each point corresponds to a different (x,y,z) position. Solutions with similar misfit occur over a wide range of depths.




**b**. Depth-misfit relationship for the inversion of a 4.2" mortar, where the initial and final ranges of models with similar misfit is tight, indicating a well defined solution.

**Figure 2.** The depth-misfit relationship, an indirect indicator of the of the depth-size ambiguity for a buried object.

*Feature Extraction: Cooperative Inversion of TEM and Magnetic Data*

In cooperative inversion, multiple data are inverted sequentially with the results of the first inversion used to constrain the second. This prior information can be formally introduced into the Bayesian formulation through the prior $p(\mathbf{m})$. Commonly utilized priors include Gaussian priors and uniform priors (i.e. a constant pdf for a parameter between two limits, and zero probability outside these limits). The solution to the inverse problem that utilizes these priors is:

$$\text{minimize } \phi(\mathbf{m}) = \sum_j \frac{1}{2\sigma_j^2}\left(m_j - \overline{m}_j\right)^2 + \frac{1}{2}\left\|V_d^{-1/2}\left(\mathbf{d^{obs}} - F(\mathbf{m})\right)\right\|^2, \text{ subject to } m_i^L \le m_i \le m_i^U \tag{13}$$

where *j* represents the index of parameters whose Gaussian pdf's are assumed to be known. The strategy we used here for cooperatively inverting magnetics and electromagnetics data, was as follows:

1) The magnetics data were inverted for a best fit dipole.
2) The dipole location was used to define $\overline{m}_j$ (for $j$ = 1, 2 and 3 which corresponds to the Easting, Northing and depth of the dipole) and the standard deviation of the parameter uncertainties used to define $\sigma_j$. The estimated model parameter standard deviations were obtained from the Gauss-Newton approximation to the Hessian at the optimum model location (e.g. Billings *et al*., 2002b).
3) The EM data were inverted using the prior obtained from the magnetics data in step 2.

Inevitably, there were anomalies in the TEM data that did not have corresponding magnetic fits and vice versa. Where no constraints from magnetometer data were available, the TEM data were inverted using the same procedure as for single inversion.

**Quality control procedures**

During the last demonstration, we visually inspected each and every two- and three-dipole inversion for the EM61 and EM63 datasets. This proved to be a time-consuming and tedious process as there were multiple views that needed to be created for each anomaly (plan-views, soundings, spatial profiles, parameters, polarization plots, termination state of optimization algorithm). In preparation for this demonstration we created some new QC views where all relevant information for each anomaly was presented on a single page and exported to a PDF document. The QC analyst scrolled through each page of the PDF and passed or failed each fit, with the results saved in UXOLab so that only the failed anomalies needed to be reinverted (see Appendix A for example views of the TEM and magnetic inversions).

### 2.1.3 Classification of Anomalies

At this stage in the process, we had feature vectors for each anomaly and needed to decide which items should be excavated as potential UXO. Rule-based classifiers use relationships derived from the underlying physics to partition the feature space. Examples include the ratio of TEM decay parameters (Pasion and Oldenburg, 2001) and magnetic remanence (Billings, 2004). For this demonstration, we focused on statistical classification techniques which have proven to be very effective at discrimination at various test-sites (e.g. Zhang *et al.*, 2003b).

Statistical classifiers have been applied to a wide variety of pattern recognition problems, including optical character recognition, bioinformatics and UXO discrimination. Within this field there is an important dichotomy between *supervised* and *unsupervised* classification. Supervised classification makes classification decisions for a *test* set comprised of unlabelled feature vectors. The classifier performance is optimized using a *training* data set for which labels are known. In unsupervised classification there is only a test set; labels are unknown for all feature vectors. Most applications of statistical classification algorithms to UXO discrimination have used supervised classification; the training data set is generated as targets are excavated. More recently, unsupervised methods have been used to generate a training data set which is an

informative sample of the test data (Carin *et al*., 2004). In addition, *semi-supervised* classifiers, which exploit both labeled data and the topology of unlabelled data, have been applied to UXO discrimination in one study (Carin *et al*., 2004).

Figure 3 summarizes the supervised classification process within the statistical framework. Given test and training data sets, we extract features from the data, select a relevant subset of these features and optimize the classifier using the available training data. Because the predicted performance of the classifier is dependent upon the feature space, the learning stage can involve further experimentation with feature extraction and selection before adequate performance is achieved.

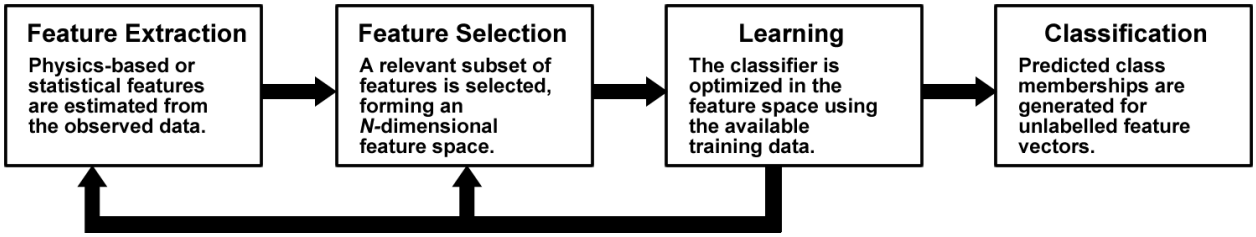



**Figure 3.** A framework for statistical pattern recognition**.**

There are two (sometimes equivalent) approaches to partitioning the feature space. The *generative* approach models the underlying probability distributions which are assumed to have produced the observed feature data. The starting point for any generative classifier is Bayes rule:

$$\mathrm{P}(\omega_i|\mathbf{x}) \propto \mathrm{P}(\mathbf{x}|\omega_i)\mathrm{P}(\omega_i) \quad . \tag{14}$$

The likelihood function $\mathrm{P}(\mathbf{x}|\omega_i)$ computes the probability of observing the feature vector **x** given the class $\omega_i$ . The prior probability $\mathrm{P}(\omega_i)$ quantifies our expectation of how likely we are to observe class $\omega_i$ . Bayes rule provides a mechanism for classifying test feature vectors: assign **x** to the class with the largest *a posteriori* probability. Contours along which the posterior probabilities are equal define decision boundaries in the feature space.

An example of a generative classifier is discriminant analysis, which assumes a Gaussian form for the likelihood function. Training this classifier involves estimating the means and covariances of each class. If equal covariances are assumed for all classes, the decision boundary is linear. While these assumptions may seem overly restrictive, in practice linear discriminant analysis performs quite well in comparison with more exotic methods and is often used as a baseline classifier when assessing performance.

Other generative classifiers assume a nonparametric form for the likelihood function. For example, the probabilistic neural network (PNN) models the likelihood for each class as a superposition of kernel functions. The kernels are centered at the training data for each class. In this case the complexity of the likelihood function (and hence the decision boundary) is governed by the width of the kernels (Figure 4).

The *discriminative* approach is not concerned with underlying distributions but rather seeks to identify decision boundaries which provide an optimal separation of classes. For example, a support vector machine (SVM) constructs a decision boundary by maximizing the *margin*

between classes. The margin is defined as the perpendicular distance between *support planes* which bound the classes, as shown in Figure 5. The decision boundary then bisects the support planes. This formulation leads to a constrained optimization problem: maximize the margin between classes subject to the constraint that the training data are classified correctly. An advantage of the SVM method over other discriminative classifiers (e.g. neural networks) is that there is a unique solution to the optimization problem.

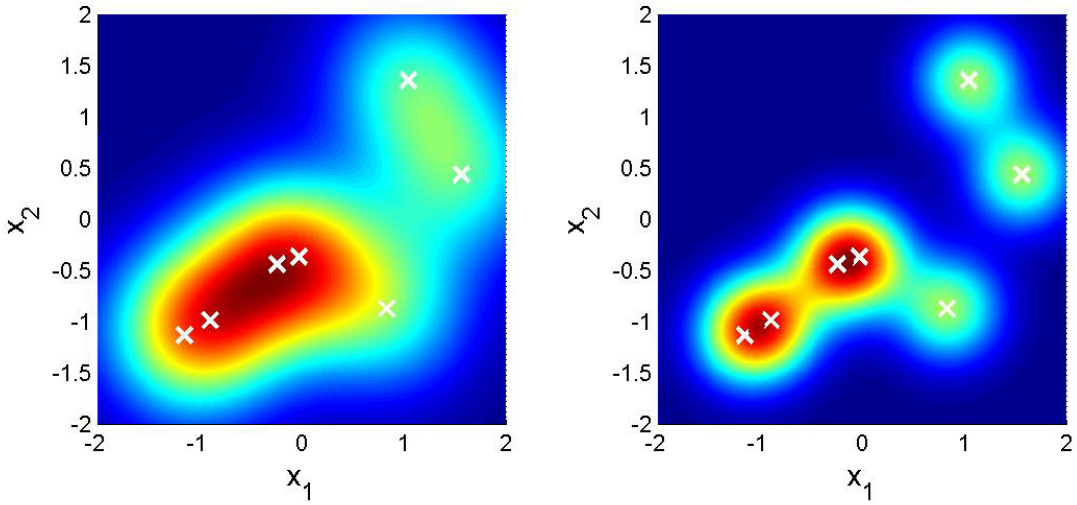



**Figure 4.** Nonparametric density estimate using Gaussian kernels. Kernel centers are shown as crosses. A large kernel width produces a smooth distribution (left) compared to a small kernel width (right).

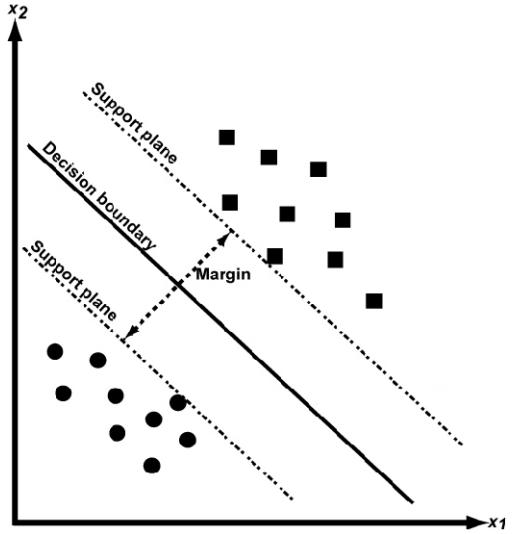



**Figure 5.** Support vector machine formulation for constructing a decision boundary. The decision boundary bisects support planes bounding the classes.

With all classification algorithms a balance must be struck between obtaining good performance on the training data and generalizing to a test data set. An algorithm which classifies all training data correctly may produce an overly complex decision boundary which may not perform well on the test data. In the literature this is referred to as "bias-variance trade-off" and is addressed by constraining the complexity of the decision boundary (regularization). In cases such as linear

discriminant analysis, the regularization is implicit in specification of the likelihood function. Alternatively, the complexity of the fit can be explicitly governed by regularization parameters (e.g. the width of kernels in a PNN or Lagrange multipliers in a SVM). These parameters are typically estimated from the training data using *cross-validation*, which sets aside a portion of the training data to assess classifier performance for a given regularization.

We obtained training data from the geophysical prove-out (GPO) and from the release of data over about 200 anomalies on the live-site.

### 2.1.4 UXOLab Software

The methodologies for data processing, feature extraction and statistical classification described above have been implemented within the UXOLab software environment. This is a Matlab based software package developed over a six year period at the UBC-GIF, principally through funding by the United States Army Corps of Engineers-Engineering Research and Development Center (USACE ERDC) (DAAD19-00-1-0120). Over the past three-years, Sky Research and UBC-GIF have considerably expanded the capabilities of the software. This is the software used for this demonstration.

## 2.2 Previous Testing of the Technology

Table 1 provides a list of some of the previous testing of the underlying methodology.

**Table 1.** Previous Inversion/Classification Testing

| Inversion/Classification Test and Location | Description | Results |
|---|---|---|
| Proof-of-concept of cooperative inversion, Yuma Proving Ground (YPG) | Test of cooperative inversion on EM63 and magnetometer data collected in 2003. TEM inversions used two decaying orthogonal dipoles, constrained using magnetics data. Three different classifiers (linear and quadratic discriminant analysis, and probabilistic neural network) were applied to the cooperative inversion results. | Classification of cooperatively inverted data is easier than inversion w/o magnetic constraints. Cleaner separation of classes is achieved for *k* parameters recovered from cooperative inversion; single and cooperative inversion results are similar for $\beta$ parameters. This test demonstrated the UXOLab capability to perform both cooperative inversion and statistical classification. |
| Geocenters STOLS EM61 and magnetometer data at Aberdeen Proving Ground (APG)/YPG | Discrimination ability of the system was marginal due to limitations in positional accuracy (5-10 cm) which is inadequate for advanced discrimination); lack of sensor orientation data; low | Results contributed to system enhancements to SKY sensor systems, including use of robotic total station (RTS) for positioning and inertial measurement unit (IMU) for sensor orientation. Demonstrated the feasibility of cooperative inversion of large volumes |

| Inversion/Classification Test and Location | Description | Results |
|---|---|---|
| | SNR; no statistical classification algorithms were applied. | of data with UXOLab. |
| Geonics EM61 and EM63 single inversion at the Rocket Range and 20 mm Range Fan at the Former Lowry Bombing and Gunnery Range. Both EM systems trialed were positioned by a Leica TPS 1206 Robotic Total Station, with orientation information provided by a Crossbow AHRS 400 IMU. The objectives of the RR surveys (8 acres) were the discrimination of a mixed range of projectiles with minimum diameter of 37 mm from shrapnel, junk, 20 mm projectiles and small-arms. The 20 mm Range Fan survey (2 acres) presented a small-item discrimination scenario where the objective was to discriminate 37 mm projectiles from ubiquitous 20 mm projectiles and 50 caliber bullets. | For the EM61, 3-dipole instantaneous amplitude models were fit to the available 4 time-channels, while for the EM63, 3-dipole Pasion-Oldenburg models were recovered from the 26 time-channel data. Parameters of the dipole model were used to guide a statistical classification. Canonical and visual analysis of feature vectors extracted form the test-plot data indicated that discrimination could best proceed using a combination of a size- and a goodness of fit-based feature vector. A SVM classifier was then implemented based on those feature vectors and using the available training data. | Two phases of digging and training were conducted at the 20mm Range Fan, and three phases at the Rocket Range (RR). At the Rocket Range, twenty-nine MK-23 practice bombs were recovered, with only one other UXO encountered (a 2.5 inch rocket warhead). At the 20 mm Range Fan, thirty-eight 37 mm projectiles (most of them emplaced) were recovered, as were a large number of 20 mm projectiles and 50 caliber bullets. For both sites, and for both instruments, the SVM classifier outperformed a ranking based on amplitude alone. In each case, the last detected UXO was ranked quite high by the SVM classifier and digging to that point would have resulted in a 60-90% reduction in the number of false-alarms. This operating point is of course unknown prior to digging. We found that using a stop-digging criteria of $f$=0 (mid-way between UXO and clutter class support planes), was too aggressive and more excavations were typically required for full recovery of detected UXO. Both the amplitude and SVM methods performed quite poorly on two deep (40 cm) emplaced 37 mm projectiles at the 20 mm Range Fan, exposing a potential weakness of the goodness of fit metric. Retrospective analysis revealed that thresholding on the size of the polarization tensor alone would have yielded good discrimination performance. |

## 2.3 Advantages and Limitations of the Technology

The main advantages of the technology are a potential reduction in the number of non-hazardous items that need to be excavated, thus reducing the costs of UXO remediation. There are two key aspects to the demonstrated technology (i) hardware and (ii) software. On the hardware side, we

are concentrating on the demonstration of commercial off-the-shelf (COTS) sensors like the EM61, EM63 and cesium (Cs) vapor magnetometer. As each of these instruments measure only one component of a vector field, a measurement at a single location provides limited information. As a consequence, relatively dense two-dimensional measurements are required for accurate recovery of relevant target parameters. These measurements must be very precisely positioned and oriented for discrimination to be successful. Strategic Environmental Research and Development Program (SERDP)/ESTCP are sponsoring the development of several next generation sensors with multi-component receivers. These newer sensors have the potential to significantly improve the estimation of target parameters using a much lower density of measurements. Over the next few years, these sensors may replace the EM61/EM63 and Cs vapor magnetometers. However, there will still be a large volume of data collected and processed with the older sensors, and there is no guarantee that any of the new sensors will be rugged and flexible enough for the diverse environments of the many hundreds of munitions and explosives of concern (MEC) contaminated sites in the country.

On the software side, advantages of UXOLab and the algorithms within the package include:

- The software contains all the functionality required to process raw geophysical data, detect anomalous regions, and perform geophysical inversion and discrimination.
- UXOLab contains algorithms for inverting magnetic and TEM datasets both separately and cooperatively using a number of different polarization tensor formulations.
- Has an extensive set of algorithms for rule-based and statistical classification algorithms.
- UXOLab has been configured in a modular fashion, so that as new sensor technologies come on-line (e.g. new TEM systems with multi-component receivers etc), the inversion functionality will be immediately available to those new sensor systems.

While UXOLab is available under license from the University of British Columbia it is not suitable for general distribution to government contractors. Firstly, using the software successfully requires advanced knowledge of geophysical inversion and statistical classification. Secondly, while the software doesn't require the user to have a Matlab license it was built entirely within the Matlab software environment to support the needs of UXO researchers. Thirdly, UBC is not set-up to provide maintenance and support for the software.

# 3 DEMONSTRATION DESIGN

This demonstration used data collected using various platforms by different demonstrators. The specific data modeling activities conducted were:

1) Dipole fitting of the MTADS magnetometer data and calculation of the magnetic remanence metric.
2) Fitting of 3-dipole beta models to the MTADS EM61 data. This hybrid model fitting approach in an attempt to prevent UXO being incorrectly modeled with 3 distinct polarizations, while allowing the model enough flexibility to model irregularly shaped shrapnel.
3) Fitting of 3-dipole beta models to the Contractor EM61 data.
4) Fitting of 3-dipole Pasion-Oldenburg models to the EM63 cued-interrogation data; and
5) Cooperative inversion of the EM61 and EM63 data using the dipole fits from the magnetometer data to constrain the object's location and depth.

For interpretation we submitted two different types of dig-sheets. The first was based on size parameters alone, while the second used all parameters and statistical classification. The following four size based interpretations were delivered:

1) Magnetics size-based (moment): A dig-sheet ranked according to decreasing size of the recovered dipole moment;
2) Magnetics size-based (remanence): A dig-sheet ranked according to magnetic remanence calculated using models designed to represent the induced magnetization of the items expected at the site;
3) MTADS EM61 size-based: Production of a dig-sheet ranked according to size (using the sum of the beta parameters for either time-channel 1).
4) Contractor EM61 size-based: The same as 2) but with the contractor dataset.

For the statistical classification, we used data over the GPO and the initial training grids to determine the feature vectors and statistical classifier to use. The following six statistically based dig-sheets were produced:

1) MTADS EM61 statistical: Using statistical classification of features derived from the MTADS EM61 data;
2) Contractor EM61 statistical: Using statistical classification of features derived from the Contractor's EM61 data;
3) EM63 statistical: Same as b) but with the EM63;
4) MTADS EM61 and magnetics statistical: As per b) but with MTADS EM61 fits constrained by the magnetics data and with the addition of the features from the magnetometer data (remanence, moment etc); and

5) EM63 and magnetics statistical: As per d) but with the EM63.

### 3.1 Performance Objectives

Table 2 lists all the performance objectives we have established for this demonstration. These performance objectives are relevant for all sensors and sensor combinations. The Institute for Defense Analysis conducted the performance confirmation.

**Table 2.** Performance Objectives

| Type of Performance Objective | Primary Performance Criteria | Expected Performance (Metric) | Actual Performance (Objective Met?) |
|---|---|---|---|
| **Quantitative** | Probability of Discrimination (PDisc) on recovered items at selected operating point | > 0.95 | Not Applicable (NA) as all 4.2" mortars were seeded |
| | PDisc on emplaced items at selected operating point | > 0.95 | Yes for all technologies |
| | False-alarm rate with PDisc (recovered) = 0.95 | > 50% reduction in false-alarms | Yes for all technologies |
| | False-alarm rate with PDisc = 1 | > 25% reduction in false-alarms | Yes for all technologies |
| | Location Accuracy of interpreted items | <0.2 m | Yes for all technologies |
| | Processing Time (interpretation) | < 5 minutes operator time per anomaly | Yes for all technologies |
| | Accuracy of inversion parameters | Within class variance of cooperative inversion < single inversion | Yes for both MTADS EM61 and EM63 |

### 3.2 Selecting Test Site(s)

The Camp Sibert ESTCP UXO Discrimination Study Demonstration site is located within the boundaries of Site 18 of the former Camp Sibert Formerly Used Defense Site (FUDS). The land is under private ownership and is used as a hunting camp.

### 3.3 Test Site History/Characteristics

Information on the Camp Sibert FUDS is available in the archival literature such as an Archives Search Report (ASR) developed in 1993. The former Camp Sibert is located in the Canoe Creek Valley between Chandler Mountain and Red Mountain to the northwest, and Dunaway Mountain and Canoe Creek Mountain to the southeast. Camp Sibert is comprised of mainly sparsely inhabited farmland and woodland and encompasses approximately 37,035 acres. The City of Gadsden is growing towards the former camp boundaries from the north. The Gadsden Municipal Airport occupies the former Army airfield in the northern portion of the site.

The site is located approximately 50 miles northwest of the Birmingham Regional Airport or 86 miles southeast of the Huntsville International Airport. The site is near exit 181 off of Interstate

59 in Gadsden and located approximately 8 miles southwest of the City of Gadsden, near the Gadsden Municipal Airport.

The area that would become Camp Sibert was selected in the spring of 1942 for use in the development of a Replacement Training Center (RTC) for the Army Chemical Warfare Service. The RTC was moved from Edgewood, Maryland to Alabama in the summer of 1942. In the fall of 1942, the Unit Training Center (UTC) was added as a second command. Units and individual replacements were trained in aspects of both basic military training and in the use of chemical weapons, decontamination procedures, and smoke operations from late 1942 to early 1945. Mustard, phosgene, and possibly other agents were used in the training. This facility provided a previously unavailable opportunity for large scale training with chemical agent. Conventional weapons training was also conducted with several types and calibers fired, with the 4.2-inch mortar being the heavy weapon used most in training.

The US Army also constructed an airfield for the simulation of chemical air attacks against troops. The camp was closed at the end of the war in 1945, and the chemical school transferred to Ft. McClellan, Alabama. The U.S. Army Technical Escort Unit (TEU) undertook several cleanup operations during 1947 and 1948; however, conventional ordnance may still exist in several locations. After decontamination of various ranges and toxic areas in 1948, the land was declared excess and transferred to private and local government ownership. A number of investigations have been conducted on various areas of the former Camp Sibert from 1990 to the present. These investigations included record searches, interviews, surface assessments, geophysical surveys, and intrusive activities.

The ESTCP UXO Discrimination Study Demonstration Site is located within the confines of Site #18, Japanese Pillbox Area No. 2, of the former Camp Sibert FUDS. Simulated pillbox fortifications were attacked first with white phosphorus (WP) ammunition in the 4.2-inch chemical mortars followed by troop advance and another volley of high explosive (HE)-filled 4.2-inch mortars. Assault troops would then attack the pillboxes using machine guns, flamethrowers, and grenades. The locations of nine possible bunkers and one trench in 1943 were identified as part of the 1999 TEC investigation. There is historical evidence of intact 4.2-inch mortars and 4.2-inch mortar debris at the site. As part of the recent investigations, a geophysical survey of Site 18 has been conducted and multiple anomalies were identified. Figure 6 is the Camp Sibert site map.

### 3.4 Present Operations

The site is no longer in active use by the military. The demonstration area is owned by a single landowner who uses the area for a hunting camp. The discrimination study was conducted after the end of the hunting season between February and August 2007.

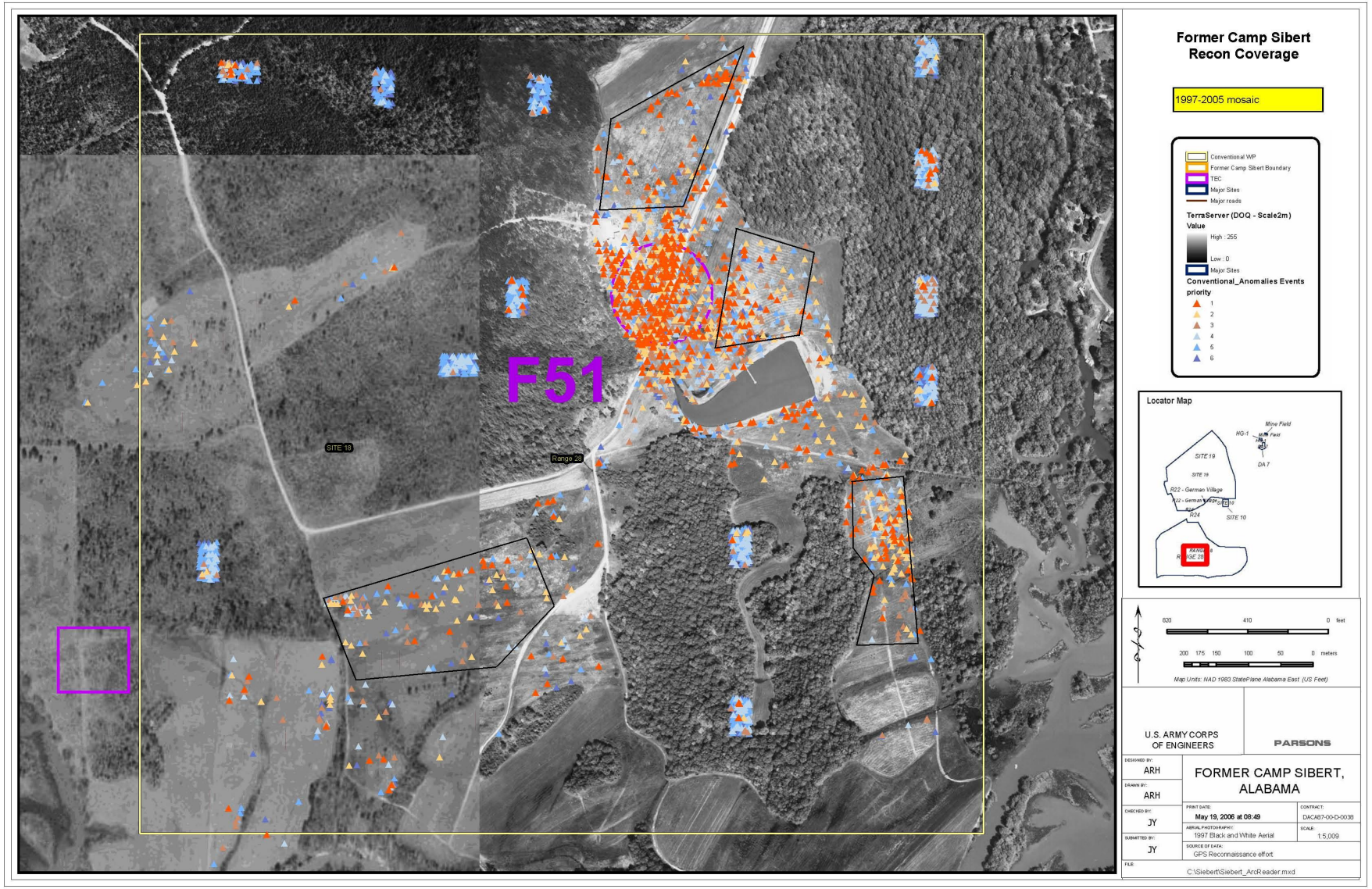



**Figure 6**. Camp Sibert Site Map.

## 3.5 Pre-Demonstration Testing and Analysis

## 3.6 Testing and Evaluation Plan

### 3.6.1 Demonstration Set-Up and Start-Up

Not applicable as this demonstration used data collected by other parties.

### 3.6.2 Period of Operation

Table 3 presents the demonstration milestones.

**Table 3.** Demonstration milestones.

| Date | Planned Action |
|---|---|
| June 15, 2007 | Receive DGM Data Reports and Target Lists for other datasets |
| July 31, 2007 | Inversions for GPO and training data complete |
| July 15, 2007 | Receive training data |
| August 1, 2007 | Submit training report to Program Office |
| September 15, 2007 | Inversions for all other grids complete |
| October 1, 2007 | Submit dig-sheets to Program Office |
| October 15, 2007 | Receive test-data from Program Office |
| November 21, 2007 | Receive scoring report from Program Office |
| February 15, 2008 | DRAFT Final Reports |
| April 15, 2008 | FINAL Reports |

### 3.6.3 Scope of Demonstration

All data modeling was performed within the UXOLab software package and included:

1) Dipole fitting of the MTADS magnetometer data and calculation of the magnetic remanence metric
2) Fitting of 3-dipole beta models to the MTADS EM61 data. As per our current demonstration, we use this hybrid model fitting approach in an attempt to prevent UXO being incorrectly modeled with 3 distinct polarizations, while allowing the model enough flexibility to model irregularly shaped shrapnel.
3) Fitting of 3-dipole beta models to the Contractor EM61 data
4) Fitting of 3-dipole Pasion-Oldenburg models to the EM63 cued-interrogation data; and
5) Cooperative inversion of the EM61 and EM63 data using the dipole fits from the magnetometer data to constrain the object's location and depth.

Different geophysicists performed modeling and QC efforts.

In section 3.1 we described the 9 different dig-sheets we submitted (four size-based, five statistical classification with two of those involving cooperative inversion). To maintain independence of each dataset we used different analysts for the different dig-sheets. The same experienced person reviewed all classification schemes and feature vectors to ensure consistency. Where cooperative inversion was used, the analyst first did the classification with the EM data and only after that dig-sheet was prepared were they provided access to the magnetics data.

### 3.6.4 Residuals Handling (not UXO identification/discrimination)

Not applicable to this effort.

### 3.6.5 Operational Parameters for the Technology

*Initial Data Processing (Creation of a Map of Geophysical Sensor Data)*

Each demonstrator provided us with a map of the appropriate geophysical sensor data and had already selected anomalies that needed to be interpreted.

*Feature Extraction*

The following steps were used to extract features over each anomaly:

- *Formation of covariance matrix (EM61 and EM63):* The base-level noise in the data (as a function of both space and time) was determined using an automated procedure, and that noise-floor was assigned to each anomaly. Together with a percentage error term the base-line error was used to form the data-covariance matrix.
- *Region definition*: For each picked anomaly, a region of data for submission to the inversion algorithm was automatically selected. Where necessary, this automated selection was manually modified by an analyst.
- *Single inversion (magnetics):* A static dipole together with a constant shift was fit to each anomaly.
- *Single inversion (EM61)*: Three-dipole instantaneous polarization models were fit to each EM61 anomaly.
- *Single inversion (EM63):* Pasion-Oldenburg parameterized three-dipole models were fit to each anomaly.
- *Cooperative inversion (EM61 and EM63):* The position and depth of the static magnetic dipole were used as constraints in the EM61 and EM63 inversions.

*Statistical Classification*

The following general procedure was used for the feature vectors of each sensor combination:

- *Selection of features:* By analysis of the training data, those features that contribute to separation of the different classes (comprising UXO types and clutter) were selected. For both the EM61 and EM63 a combination of a size and time-decay feature vector provided reliable classification ability.

- *Choice of classification algorithm:* Through analysis of the training data the best performing classifier(s) was selected: a Probablistic Neural Network.
- *Classification:* Anomalies were placed on a prioritized dig-list by using the classifier to compute probabilities of class membership for unlabeled feature vectors. The probability of membership of the UXO class was reported on the dig-sheet.
- *Anomalies where feature vectors are unreliable:* Some anomalies had insufficient SNR or data coverage to constrain the TEM model parameters. This included anomalies with overlapping signatures that could not be isolated and inverted one at a time. All these anomalies were placed in the dig-sheet and given a label of "can't analyze" and were excavated as suspected UXO.

*Memo for Program Office once training data are analyzed*

Following analysis of the GPO and training data we submitted a Memo to the Program Office that:

- Described the feature vectors that were estimated, those that were selected for input to the classifier, and why;
- Described how dig-sheet thresholds were specified;
- Summarized the analysis of GPO data to estimate target parameters, including documentation of the variability and accuracy in key parameters.

This memo is reproduced as Appendix B of this demonstration report.

#### 3.6.6 Demobilization

Not required. This effort relied on data provided by other demonstrators.

#### 3.6.7 Health and Safety Plan (HASP)

Not applicable to this effort.

### 3.7 Selection of Analytical/Testing Methods

Not applicable to this effort

### 3.8 Selection of Analytical/Testing Laboratory

Not applicable to this effort

### 3.9 Management and Staffing

Table 4 lists the responsibilities of staff involved in the demonstration.

**Table 4**. Responsibilities of staff for the demonstration.

| Personnel | Title | Responsibilities |
|---|---|---|
| Stephen Billings (Sky Research) | Principal Investigator (PI) and QC magnetics | Technical oversight of entire project and QC of all magnetic inversions and the EM61 contractor data |
| Joy Rogalla (Sky Research) | Project manager | Manage resources for the project, oversee the cost-tracking for the demonstration |
| Len Pasion (Sky Research) | QC feature vectors | Provide technical oversight and QC of all EM inversions |
| Laurens Beran (UBC-GIF) | QC statistical classification | Provide technical oversight and QC of all statistical classifications |
| Jon Jacobson (Sky Research) | Analyst | Perform single and cooperative inversions |
| Kevin Kingdon (Sky Research) | Analyst | Perform single and cooperative inversions |
| David Sinex (UBC-GIF) | Analyst | Implement any changes required to UXOLab |
| LinPing Song (UBC-GIF) | Analyst | Perform single and cooperative inversions |
| Nicolas Lhomme (UBC-GIF) | Analyst | Perform single and cooperative inversions |

In Table 5 we list the inversion responsibilities of each of the four analysts that were involved in the project. Data were processed in blocks corresponding the South-East 1 (SE1), South-East 2 (SE2), South-West (SW) and GPO areas of the project site. The SW portion of the site was quite large and was subsequently split into two pieces: SW1 and SW2.

**Table 5.** Block locations of inversions for each dataset performed by each analyst.

| Person | Magnetics | MTADS EM61 | EM61 Contractor | EM63 | MTADS EM61 cooperative | EM63 cooperative |
|---|---|---|---|---|---|---|
| **Jacobson** | SE1, SW2 | SW1, SE2 | | | SW1, SE2 | |
| **Kingdon** | SW1, SE2 | SE1, SW2 | | ALL | SE1, SW2 | ALL |
| **Song** | | | SE1, SW2 | | | |
| **Lhomme** | | | SW1, SE2 | | | SE1, SE2 |
| **Quality control** | Billings | Lhomme/ Pasion | Billings | Pasion | Lhomme/ Pasion | Pasion |

To maintain independence of each dataset we had intended to use different analysts to produce the different dig-sheets. However, we decided to use the same feature vectors and statistical classification for each of the TEM methods, and therefore all the dig-sheets were produced by the PI.

# 4 PERFORMANCE ASSESSMENT

## 4.1 Performance Criteria

The effectiveness of the demonstration was evaluated according to the performance objectives cited in Section 3.1. Table 6 summarizes these objectives and the associated metrics for evaluation

**Table 6.** Performance Criteria.

| Performance Criteria | Description | Primary or Secondary |
|---|---|---|
| Probability of Discrimination (recovered) | (# of MEC items detected and recommended for excavation) / (# MEC items detected) | *Primary* |
| Probability of Discrimination (emplaced) | (# of emplaced items detected and recommended for excavation) / (# emplaced items detected) | *Primary* |
| False alarm rate (FAR) | # of anomalies not corresponding to an ordnance item | *Primary* |
| Probability of False alarm (Pfa) | # false positives (i.e. declaration of ordnance) corresponding to clutter/# of opportunities for false positive | *Primary* |
| Geo-reference position accuracy | Distance to interpreted items | *Primary* |
| Processing Time (Interpretation) | Total minutes of operator time per anomaly | *Secondary* |
| Accuracy of inversion parameters | Comparison of spread in parameters for a given ordnance class for cooperative versus single inversion | *Secondary* |

## 4.2 Performance Confirmation Methods

Table 7 lists the specific performance confirmation methods for each metric.

**Table 7.** Expected Performance and Performance Confirmation Methods. The geo-reference metric comprises the percentage of items within 20 cm of the ground-truth location. Metrics that were successfully met are shown in green, while those that were not are shown in yellow.

| Performance Criteria | Expected Performance Metric | Performance Confirmation Method | Magnet-meter moment | Magnet-meter remanent | EM61 size | EM61 classifier | MTADS EM61 size | MTADS EM61 classifier | MTADS EM61 coop. | EM63 | EM63 coop. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pdisc (recovered) at operating point | ≥ 0.95 | Compare to ground-truth | 1.00 | 0.98 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| False-alarm rate at operating point | Not specified | Compare to ground-truth | 0.28 | 0.24 | 0.67 | 0.62 | 0.58 | 0.56 | 0.26 | 0.33 | 0.18 |
| False-alarm rate with PDisc = 1 | FA reduced > 25% | Compare to ground-truth | 0.24 | 0.31 | 0.48 | 0.41 | 0.55 | 0.48 | 0.28 | 0.28 | 0.14 |
| Geo-reference position accuracy | < 0.2 m | Compare to ground-truth | 86% | 86% | 61% | 61% | 66% | 66% | 82% | 86% | 89% |
| Processing Time (interpretation) | < 5 minutes per anomaly | Data analysis log | 3.7 | 3.7 | 9.1 | 9.1 | 5.0 | 5.0 | 7.4 | 5.3 | 8.8 |
| Inversion accuracy | Cooperative better than single | Compare to ground-truth | NA | NA | NA | NA | NA | NA | Yes for size | NA | Yes |

Performance confirmation was conducted by the Institute for Defense Analysis (IDA). The rest of this section is extracted from a preliminary IDA document on the scoring protocols.

- For each of the discrimination algorithms to be scored, the demonstrator submitted a ranked dig list as described in Figure 7:
- The first item in the list (Rank = 1) should be that which you are most certain does NOT need to be dug up (shown in green).
- The bottom items should be those that you are most certain are munitions and must be dug (shown in red). Thus, larger numerical rankings are associated with likely targets of interest.
- A threshold should be set at the point beyond which you would recommend digging all targets, either because you are certain they are ordnance or because a high confidence determination cannot be made (heavy black dividing line in Figure 7).
- Two other bands should be specified indicating (1) the range of targets where the SNR, data quality or other factors prevent any meaningful analysis (shown in grey), and (2) the range of targets where the data can be fit in a meaningful way, but the derived parameters do not permit a conclusion (shown in yellow). These represent two levels of "guessing."

| Rank | Comment |
|---|---|
| 1 | |
| 2 | High confidence NOT ordnance (no dig) |
| 3 | |
| … | Can't make a decision (dig) |
| … | |
| … | |
| … | Can't analyze (dig) |
| … | |
| … | Can't make a decision (dig) |
| … | |
| 97 | |
| . | High confidence ordnance (dig) |
| . | |
| N | |

**Figure 7.** Ranked dig list. High numbers represent likely UXO.

The ranked list contained the anomaly ID from the sensor *Associated Detection List*, anomaly depth, and values for characterization parameters used that are appropriate to the algorithm (e.g., betas, decay constants, remnant magnetization measures, size, etc.).

The following ground rules were be followed in discrimination scoring

- Each algorithm was scored for discrimination against all dug anomaly calls that appear on the *Associated Detection List* of the sensor data employed by the algorithm. The discrimination score was not penalized by items the sensor did not detect.

- Any joint or cooperative discrimination algorithms were scored on the union of the detection calls of sensors employed.

The following metrics were provided by IDA based on discrimination scoring of the ranked dig list for each of the discrimination algorithms:

- Pdisc (emplaced) = (Number of emplaced UXO whose rank number is greater than the recommended dig threshold)/(Number of emplaced UXO detected by sensor used for discrimination)
- Pdisc (recovered) = (Number of recovered targets of interest (including emplaced targets) whose rank number is greater than the recommended dig threshold)/(Number of recovered targets of interest (including emplaced targets) detected by sensor used for discrimination)
- Probability of False Alarm (Pfa) at dig threshold = (Number of non-targets of interest whose rank number is greater than the dig threshold)/(Total number of non-targets of interest on detection list for sensor used for discrimination)
- Pfa at 100% Pdisc = (Number of non-targets of interest whose rank number is greater than that of the lowest rank number target of interest)/(Total number of non-targets of interest on detection list for sensor used for discrimination)
- A conventional receiver operating characteristic (ROC) curve was plotted for each dig list by plotting Pdisc (recovered) vs. Pfa, using the discrimination ranking (bottom to the top of each dig list).

## 4.3 Discrimination performance of each dataset

The feature vector selection and discrimination strategy for each sensor combination were described in a memo sent to the Program Office in September 2007. This memo is reproduced in Appendix B. Discrimination results for each of the 9 sensor datasets are summarized in Table 8.

### 4.3.1 Size-based discrimination ranking (magnetometer, EM61 and MTADS EM61)

In this section we investigate the performance of the following four methodologies:

- Magnetometer data ranked by moment;
- Magnetometer data ranked by remanence;
- EM61 data ranked by size, $L_1(t_1)$; and
- MTADS EM61 data ranked by size, $L_1(t_1)$.

For magnetics, the training data revealed a number of 4.2” mortars with relatively large remanence (Figure 8a). All 4.2” mortars had moments greater than 0.17 $Am^2$, and could therefore be considered “large”. Consequently, the size of the moment (and not the remanence) was used to prioritize the dig-list. We also produced a remanence prioritized dig-list so that we could test performance using that discrimination metric.

All recovered moments from 4.2” mortars in the blind-test data were above the stop-digging threshold (or were listed as “can’t analyze”), whereas two items in the remanence ranked list lay

outside the region recommended for excavation (Figure 8b). There were 95 items in the blind-test data classified as “can’t analyze”. The ROC curves for diglists ranked by moment and remanence are shown in Figures 9 a and c (including “can’t analyze”) and 9 b and d (excluding “can’t analyze”). The last UXO occurs after 170 (75) false positives (FP) for the ranking by moment 218 (123) FP for the ranking by remanence, where the numbers is brackets exclude the “can’t analyze” category. The remanence ranking is initially more efficient at recovering UXO than the moment ranking with 80% of 4.2” mortars recovered with 15 FP compared to 45 FP. The final 20% of 4.2” mortars contain some rounds with high-remanence which causes the moment ranking to be more efficient at recovery of all rounds.

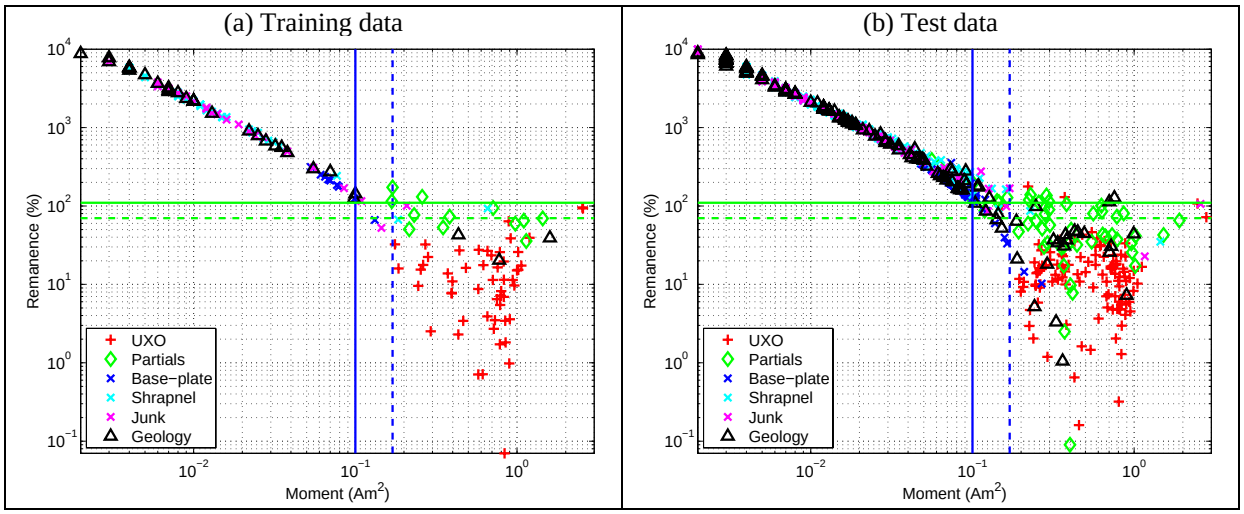
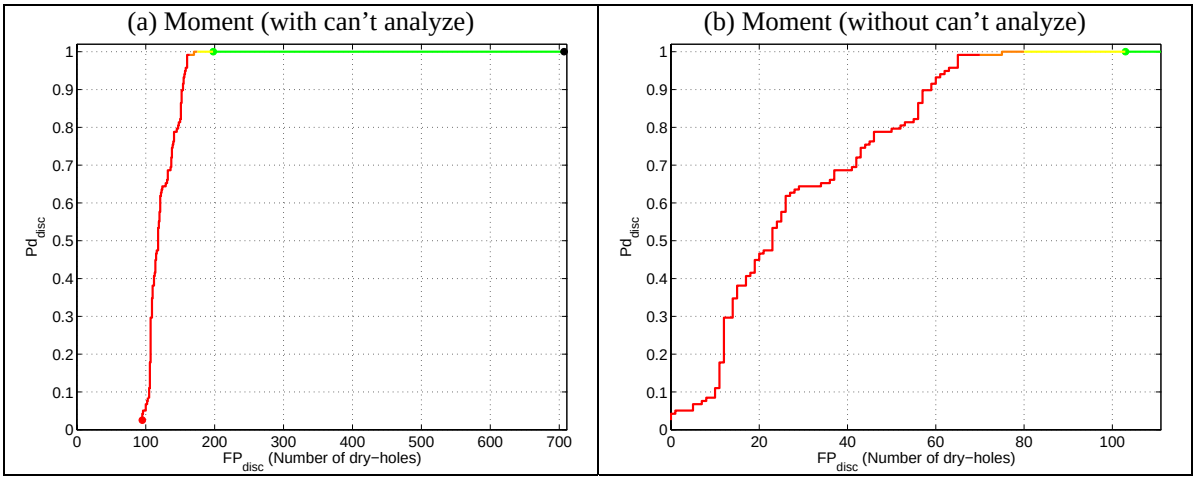



**Figure 8.** Scatterplot of the moment versus remenance for the (a) training data; and (b) test data.

The size based ranking for the EM61 and MTADS EM61 datasets were based on the magnitude of the primary polarization at the first time channel: $L_1(t_1)$. Both EM sensors generated more “can’t analyze” anomalies than the magnetics, with 149 and 285 for the EM61 and MTADS EM61 respectively (Table 8). As discussed in section 4.7.4, many of the MTADS EM61 “can’t analyze” anomalies were due to noise generated by cart-bounce. When excluding the “can’t analyze” category, both EM datasets require less FP (56 and 53 FP for the EM61 and MTADS EM61) than the magnetics to recover all 4.2” mortars (Figure 9).



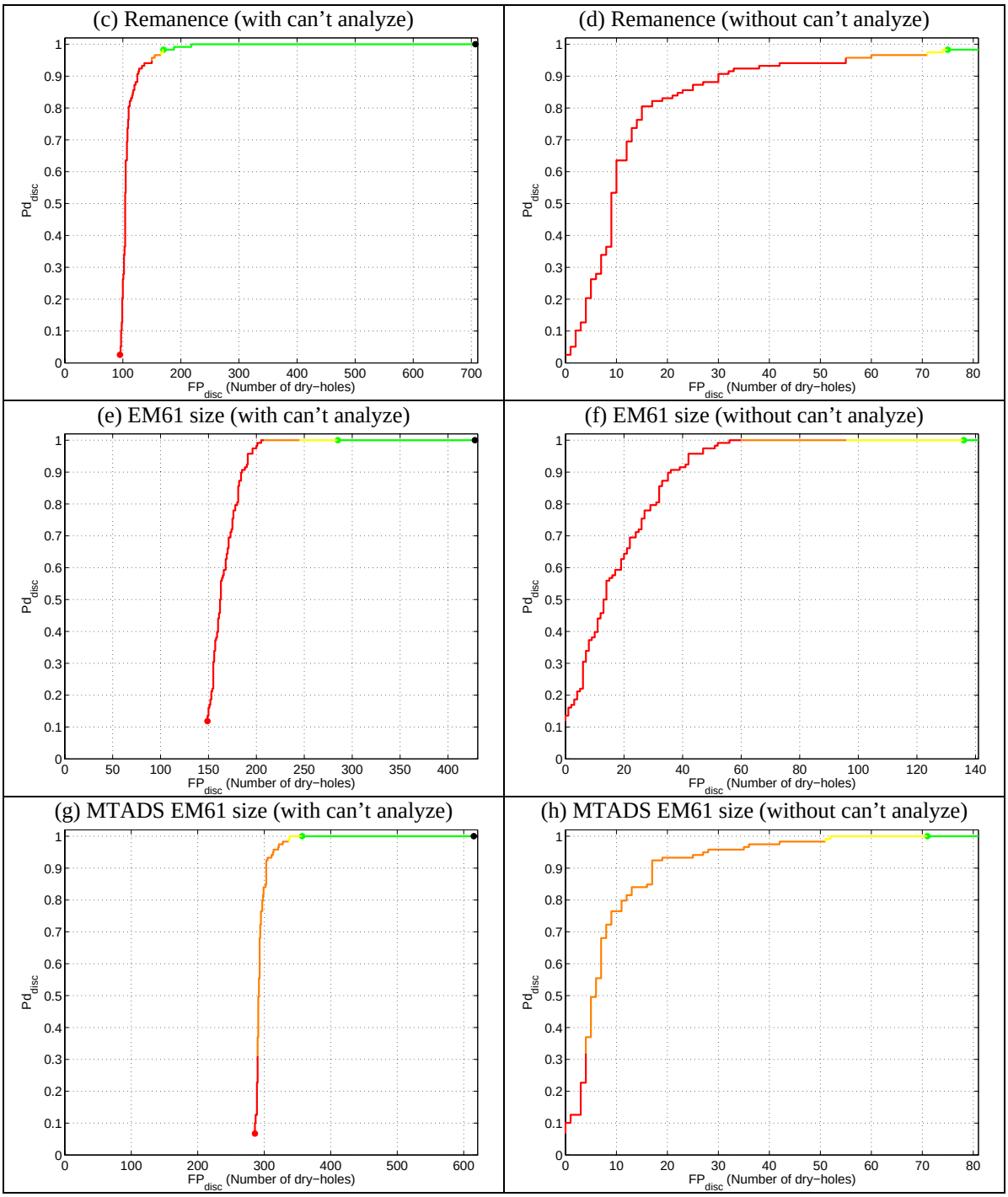



**Figure 9.** ROC curves for the contractor EM61 (a and b), MTADS EM61 (c and d) and MTADS EM61 cooperative (e and f) datasets. The plots on left include the "can't analyze" category, while those on the right exclude them.

**Table 8.** Summary of the discrimination performance of the 9 different sensor combinations or discrimination methods. Numbers in brackets represent the results excluding the "can't analyze" category.

| | | | At operating point | | | | At Pdisc = 1 | |
|---|---|---|---|---|---|---|---|---|
| **Technologies** | **# alarms** | **# UXO** | **# Can't analyze** | **# false positives** | **Pdisc** | **Pfa** | **# false alarms** | **Pfa** |
| Magnetometer (moment) | 825 | 118 | 95 (2) | 198 (103) | 1 | 0.28 | 170 (75) | 0.24 |
| Magnetometer (remanence) | 825 | 118 | 95 (2) | 170 (75) | 0.98 | 0.24 | 218 (123) | 0.31 |
| EM61 (size) | 546 | 118 | 149 (13) | 285 (136) | 1 | 0.67 | 205 (56) | 0.48 |
| EM61 (classification) | 546 | 118 | 149 (13) | 264 (115) | 1 | 0.62 | 174 (25) | 0.41 |
| MTADS EM61 (size) | 734 | 119 | 285 (8) | 357 (72) | 1 | 0.58 | 338 (53) | 0.55 |
| MTADS EM61 (classification) | 734 | 119 | 285 (8) | 344 (59) | 1 | 0.56 | 293 (8) | 0.48 |
| MTADS EM61 cooperative | 734 | 119 | 205 (2) | 275 (70) | 1 | 0.45 | 208 (3) | 0.34 |
| MTADS EM61 cooperative (with magnetics) | 983 | 119 | 226 (2) | 307 (81) | 1 | 0.26 | 240 (14) | 0.28 |
| EM63 | 150 | 34 | 30 (0) | 38 (8) | 0.97 | 0.33 | 47 (17) | 0.41 |
| EM63 cooperative | 150 | 34 | 16 (0) | 21 (5) | 1 | 0.18 | 16 (0) | 0.14 |

### 4.3.2 Ranking by classification (EM61, MTADS EM61 single and cooperative)

In this section we investigate the performance of the following three methodologies:

- Statistical classification of the EM61 data;
- Statistical classification of the MTADS EM61 data; and
- Statistical classification of the MTADS EM61 cooperatively inverted data.

As described in Appendix B, for each of these sensor combinations we used a size based and a time-decay based feature vector and a PNN classifier. Figure 10 plots the two feature vectors for each sensor combination over the decision surface: and has separate plots for the training and test-data. Feature vectors within the UXO, partial rounds and base-plate classes are more tightly clustered for the MTADS EM61 than for the contractor EM61, with further improvement evident in the MTADS EM61 cooperatively inverted. For all three datasets we used two decision boundaries: a more aggressive one for high Figure of Merit (FOM) anomalies, and a less aggressive one for low FOM anomalies. For each of the three datasets, all 4.2" mortars lie on the appropriate side of the "high-confidence UXO" decision boundary. As shown in Figure 11, all potentially hazardous items are recovered with very few false-positives (ignoring the "can't analyze" category) with FP = 25, 8 and 3 for EM61, MTADS EM61 and MTADS EM61 cooperative, respectively. The discrimination performance of both MTADS EM61 datasets are degraded by the high number of can't analyze items: 285 in the MTADS EM61, compared to 205 in the MTADS EM61 cooperative. The lower number of can't analyze items in the cooperatively inverted data occurs because the magnetics data removes some of the depth ambiguity inherent in the EM data.

Figure 12 compares the magnetometer, EM61, MTADS EM61 and MTADS EM61 cooperative ROC curves. When including the "can't analyze" category (Figure 12a), the magnetometer data requires the least number of FP excavations at its operating point (170 compared to 264, 344 and 275 for the EM61, MTADS EM61 and MTADS EM61 cooperative). 72% of the non-UXO items can be left in the ground with the magnetometer data compared to 33, 44 and 55% for the other datasets. When excluding the can't analyze category (Figure 12b), it's evident that feature vectors extracted from the MTADS EM61 and MTADS EM61 cooperatively inverted data are more highly discriminatory than the magnetometer or EM61.

The IDA demonstration protocol involved scoring the cooperatively inverted data on the union of the EM and magnetometer datasets. There were 249 magnetometer anomalies that did not have a corresponding EM anomaly. Only 32 of these required excavation using the same criterion utilized in the moment ranked digsheet, with 21 of those anomalies in the "can't analyze" category. This modest increase in the number of anomalies to excavate using cooperative inversion is more than compensated by the 80 less "can't analyze" anomalies.

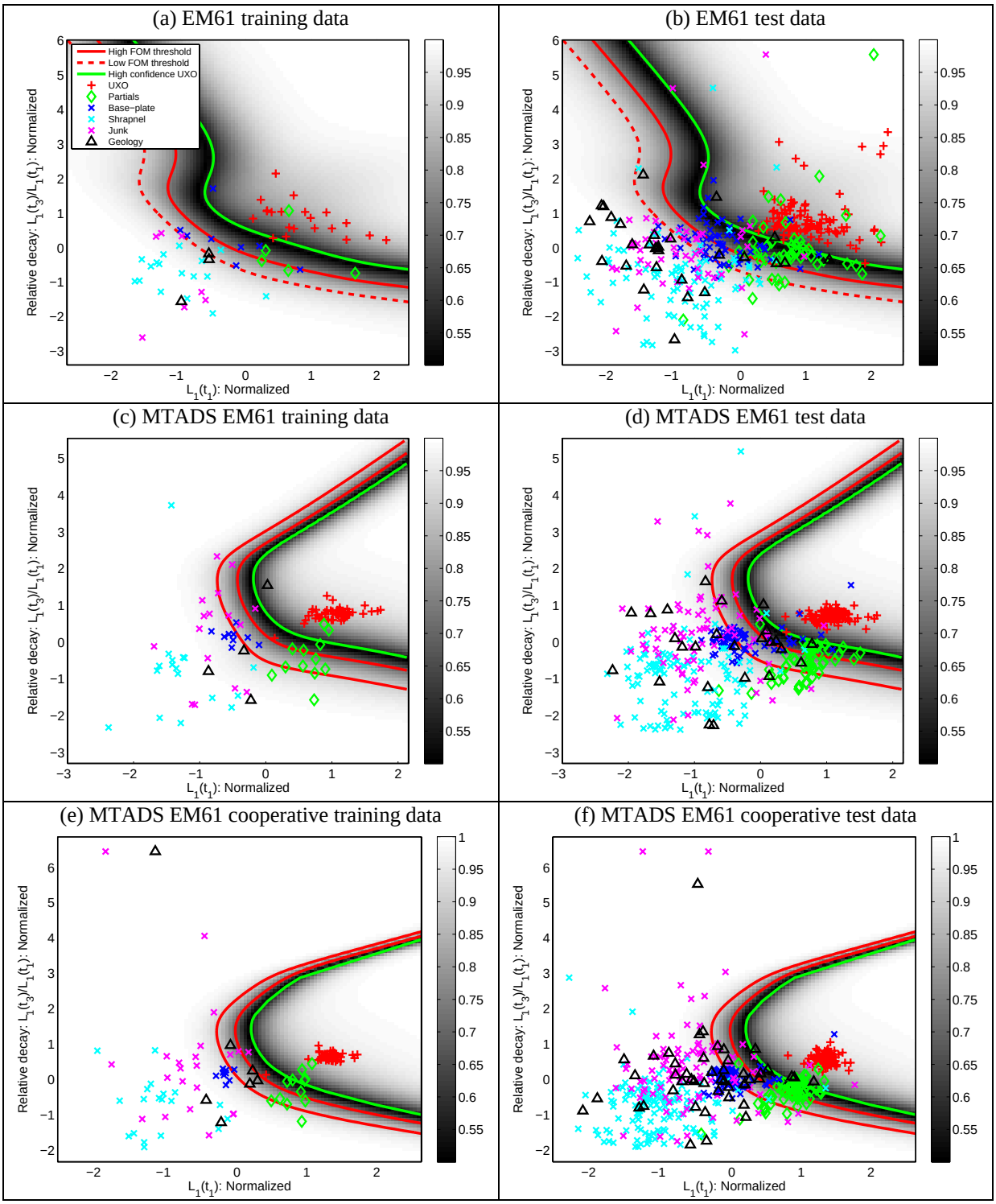



**Figure 10.** Feature vector plots for the contractor EM61 (a and b), MTADS EM61 (c and d) and MTADS EM61 cooperative (e and f) datasets. The plots on left are for the training data, while those on the right are test-data.

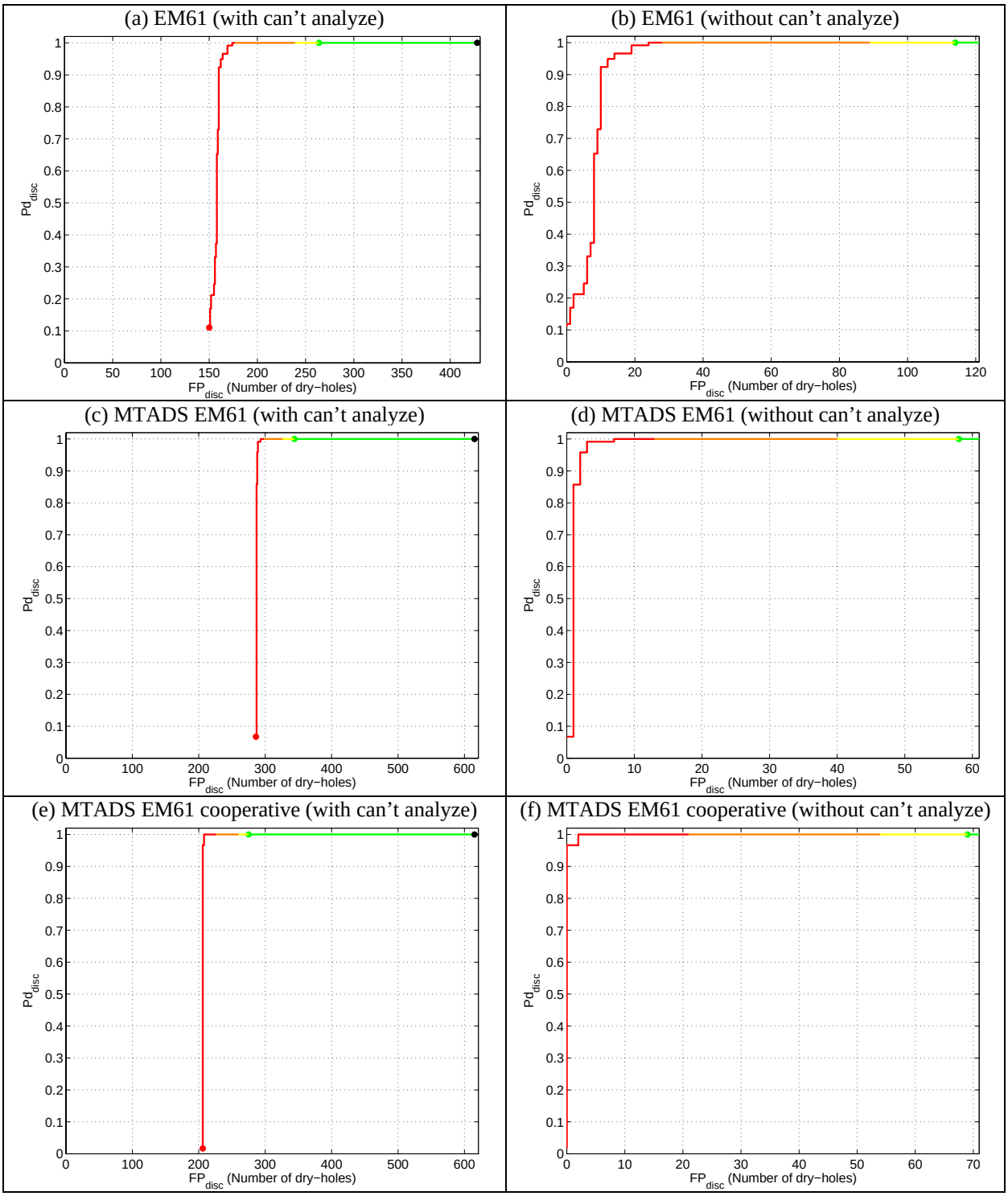



**Figure 11.** ROC curves for the contractor EM61 (a and b), MTADS EM61 (c and d) and MTADS EM61 cooperative (e and f) datasets. The plots on left include the "can't analyze" category, while those on the right exclude them.

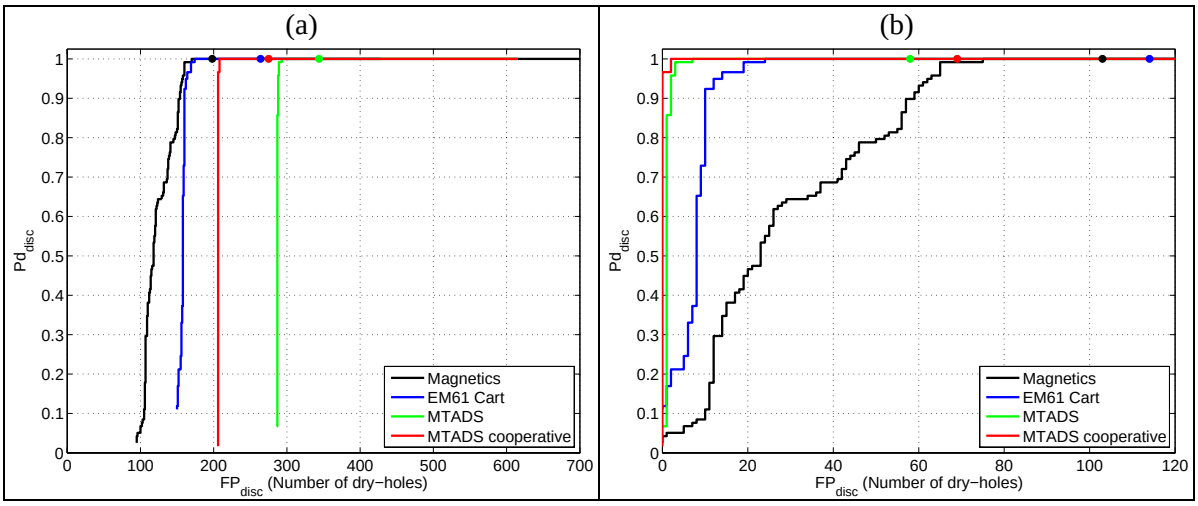



**Figure 12.** Comparision of ROC curves for the magnetics (moment), contractor EM61, MTADS EM61 and MTADS EM61 cooperative datasets: (a) including the "can't analyze" category; (b) excluding can't analyze anomalies.

### 4.3.3 Cued-interrogation data (EM63 single and cooperative)

In this section we investigate the performance of the following two cued-interrogation methodologies:

- Statistical classification of the EM63 data; and
- Statistical classification of the EM63 cooperatively inverted data.

As with the EM61 based datasets, discrimination was based on a PNN classifier with a size-based and a time-based feature vector (see Appendix B). The size-based feature vector was the $k_1$ parameter from the primary polarization with the ratio of primary polarizations at the $15^{th}$ and $1^{st}$ time channels used as the time-decay parameter. The UXO, partial rounds and base-plate classes are tightly clustered in both datasets, with less variation in the cued-interrogation data (Figure 13). The UXO class for the EM63 data contain two outliers, one of which causes the false-negative evident in Figures 14 a and c. The EM63 when cooperatively inverted produces a "perfect" ROC curve with 0 FP (excluding can't analyze anomalies) at the point where all UXO are recovered. At the operating point a total of 21 FP, with 16 of these in the "can't analyze" category, are required, with 82% of non-UXO left in the ground. In Figure 14 we also show an ROC curve for the MTADS EM61 cooperative when restricted to the same 150 cued-interrogation anomalies. It also results in a "perfect" ROC curve but requires 34 FP at the operating point, with 6 of those in the "can't analyze" category.

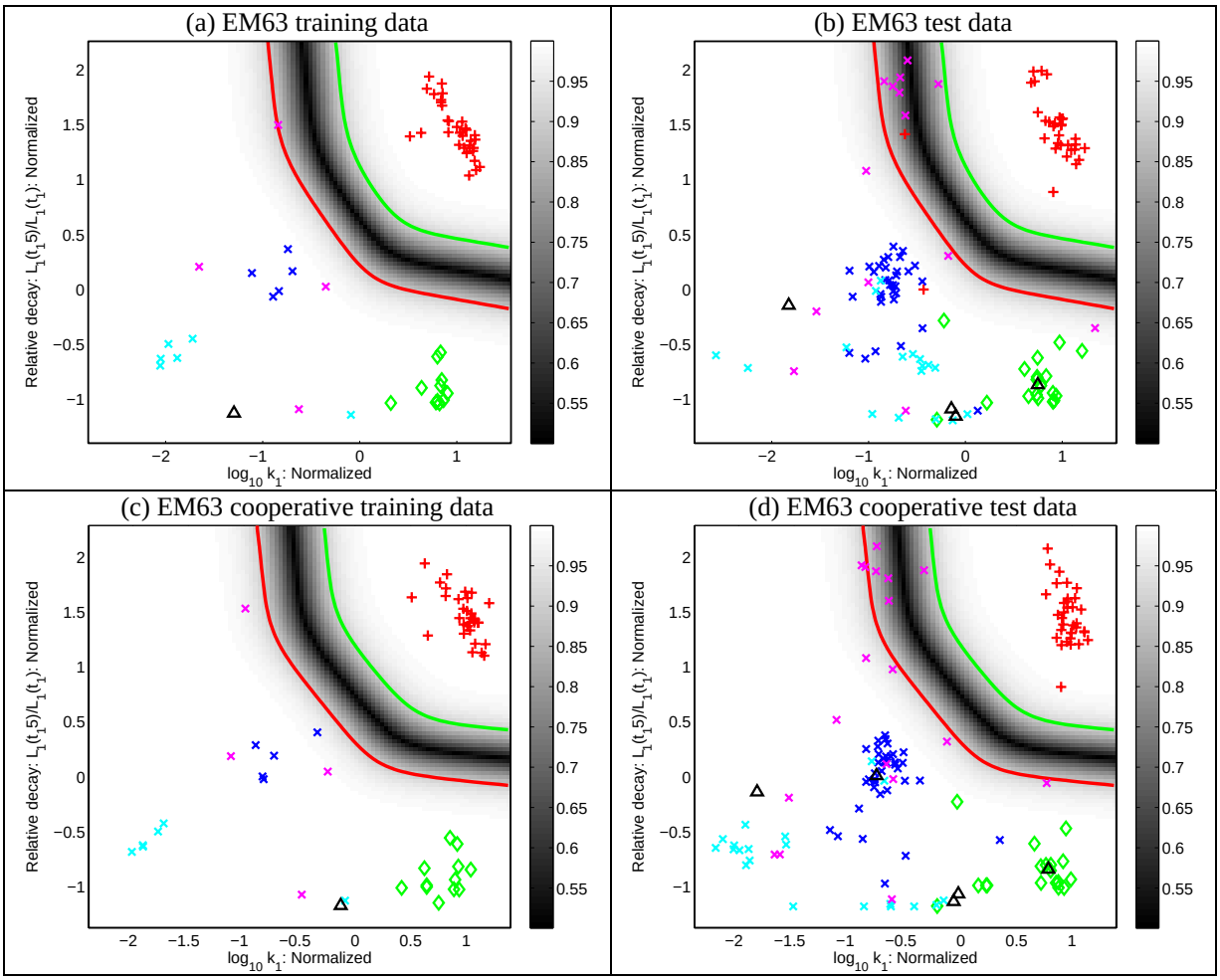



**Figure 13.** Feature vector plots for the contractor EM63 (a and b) and EM63 cooperative (c and d) datasets. The plots on left are for the training data, while those on the right are test-data.

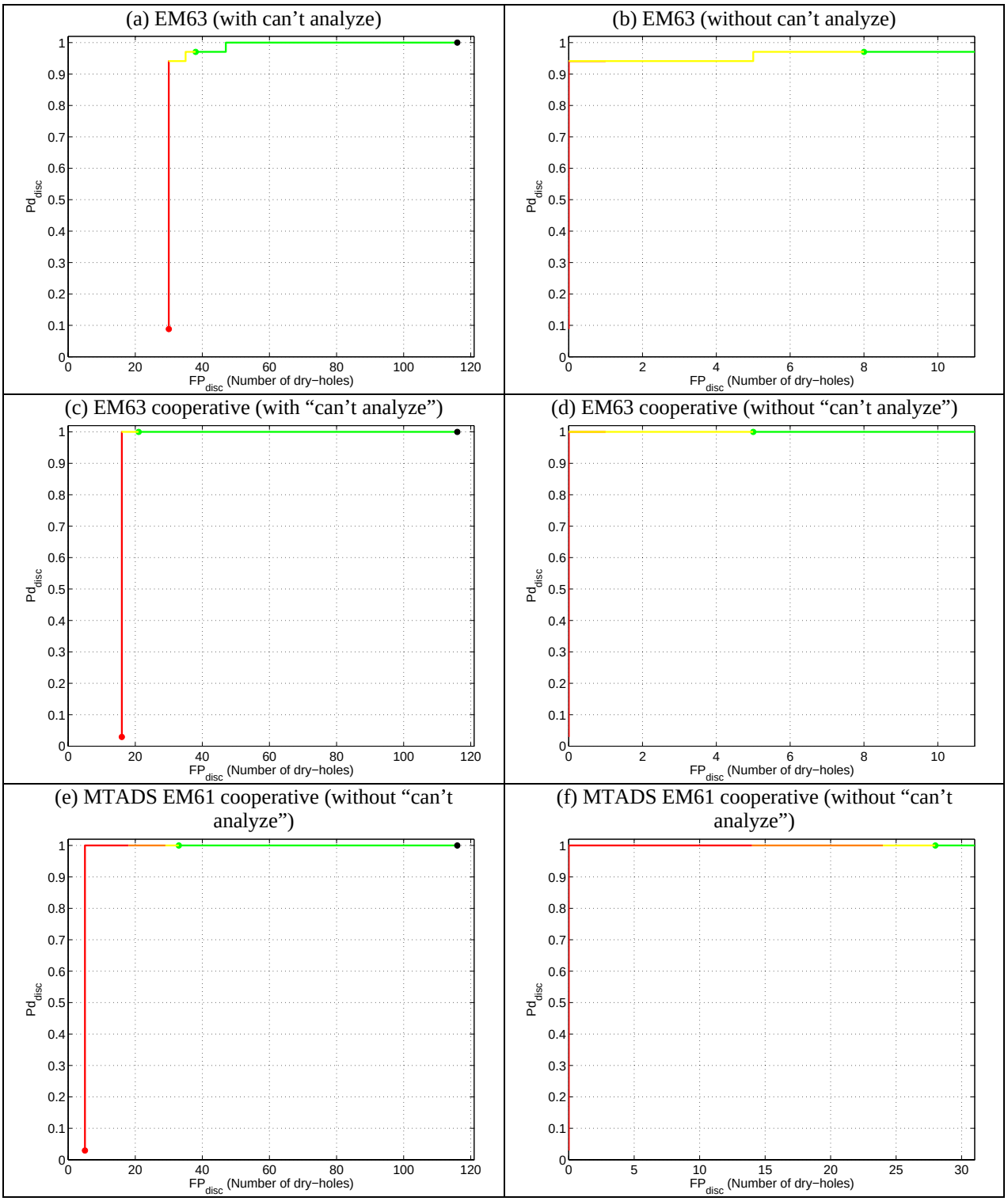



**Figure 14.** ROC curves for the EM63 (a and b), EM63 cooperative (c and d) and MTADS EM61 cooperative (e and f) on each of the cued-interrogation anomalies. The plots on left include the "can't analyze" category, while those on the right exclude them.

## 4.4 Accuracy of inverted positions and depths

The position and depth of each excavated item were carefully measured and made available to each of the demonstrators after IDA had completed their scoring analysis. The EM61 dataset exhibited a positional bias of 9.9 cm East and 24.5 cm South (indicating that the geographic positions of the data weren't adjusted to coincide with the Program Office coordinate system). None of the other datasets exhibited a bias of greater than 2.5 cm in any direction (Table 9). Figure 15 shows a scatter plot of the error in Easting and Northing for each of the six-different sensor combinations. The analysis includes the GPO, training and test data but excludes the "can't analyze" category and any items identified as geology or soils. The cumulative distribution in Figure 17 reveals that the estimated positions from the magnetometer and EM63 data tend to be better positioned than the EM61 cart or MTADS EM61 data. This is not surprising given the greater density of along-line measurements of the magnetometer (25 cm) and EM63 (30 cm) compared to the EM61 cart and MTADS EM61 (50 cm). The use of the magnetometer location constraints significantly improves the accuracy of the MTADS EM61 inverted positions.

Scatter plots of the predicted versus ground-truth depths for each of the six different sensor combinations are shown in Figure 16. There is excellent agreement between estimated and actual depths for the magnetometer with 85% within 10 cm and 96% within 20 cm (Figure 17a). The EM61 cart and MTADS EM61 display a much larger scatter in actual and predicted depths with 74 and 76% respectively within 20 cm. For the EM61, the estimated depths of the deeper 4.2" mortars are particularly poor and contribute to the relatively wide range of estimated size observed for that class. Both the EM61 cart and MTADS EM61 have a tendency to predict deeper depths for small, shallow items. This characteristic was also exhibited by the EM61 and EM63 data at FLBGR during the last demonstration conducted under this ESTCP project.

Cooperative inversion considerably improves the accuracy of the estimated depths: 75% of MTADS EM61 depths are within 10 cm compared to 53% when inverted without magnetometer depth constraints. The performance gain is from 75 to 87% within 10 cm when the EM63 data are cooperatively inverted.

**Table 9.** Position biases, distance and depth to 90$^{th}$ percentile for each sensor combination.

| Sensor | Easting bias (cm) | Northing bias (cm) | Distance 90$^{th}$ percentile (cm) | Depth 90$^{th}$ percentile (cm) |
|---|---|---|---|---|
| Magnetometer | 0.0 | 0.0 | 24 | 13 |
| EM61 | 9.9 | -24.5 | 46 | 37 |
| MTADS EM61 | -0.1 | 2.1 | 54 | 33 |
| MTADS EM61 cooperative | 0.0 | 0.0 | 31 | 19 |
| EM63 | 0.3 | -2.0 | 24 | 19 |
| EM63 cooperative | 0.0 | -2.0 | 21 | 13 |

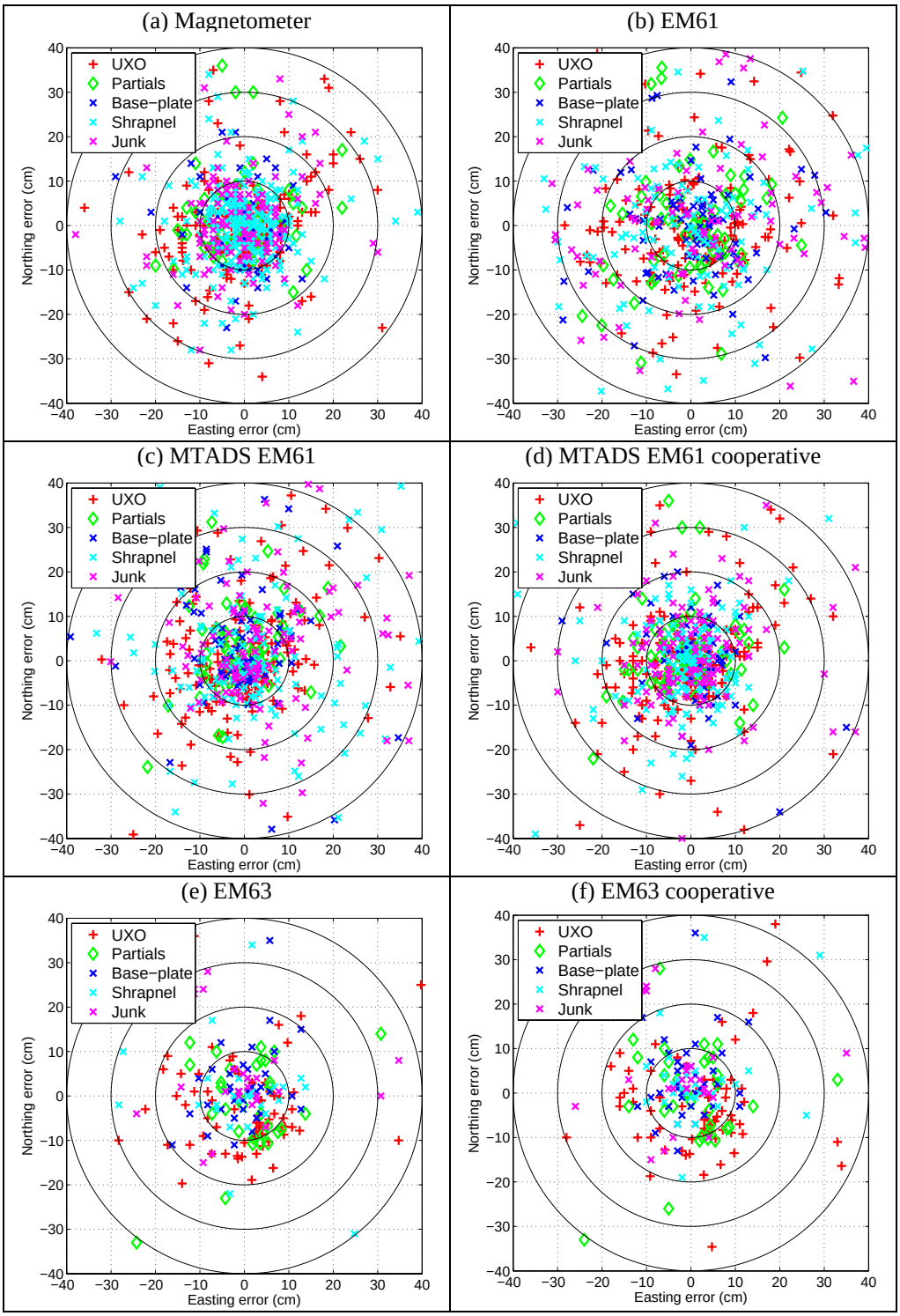



**Figure 15.** Scatter plot of fitted minus ground-truth positions for (a) magnetometer; (b) EM61; (c) MTADS EM61; (d) MTADS EM61 cooperative; (e) EM63; and (f) EM63 cooperative.

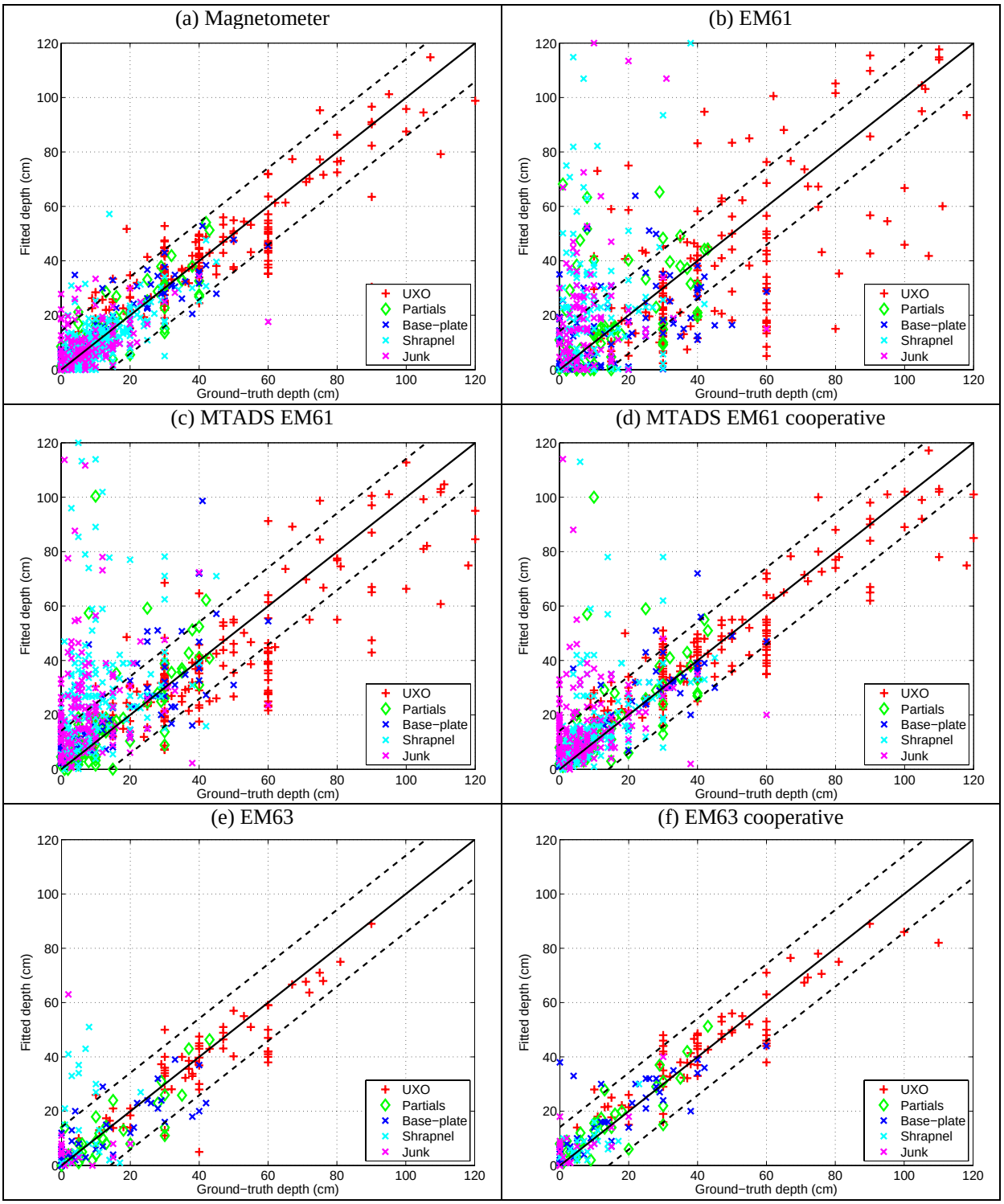



**Figure 16.** Scatter plot of fitted versus ground-truth depths for (a) magnetometer; (b) EM61; (c) MTADS EM61; (d) MTADS EM61 cooperative; (e) EM63; and (f) EM63 cooperative. The dashed lines represent errors of plus and minus 15 cm.

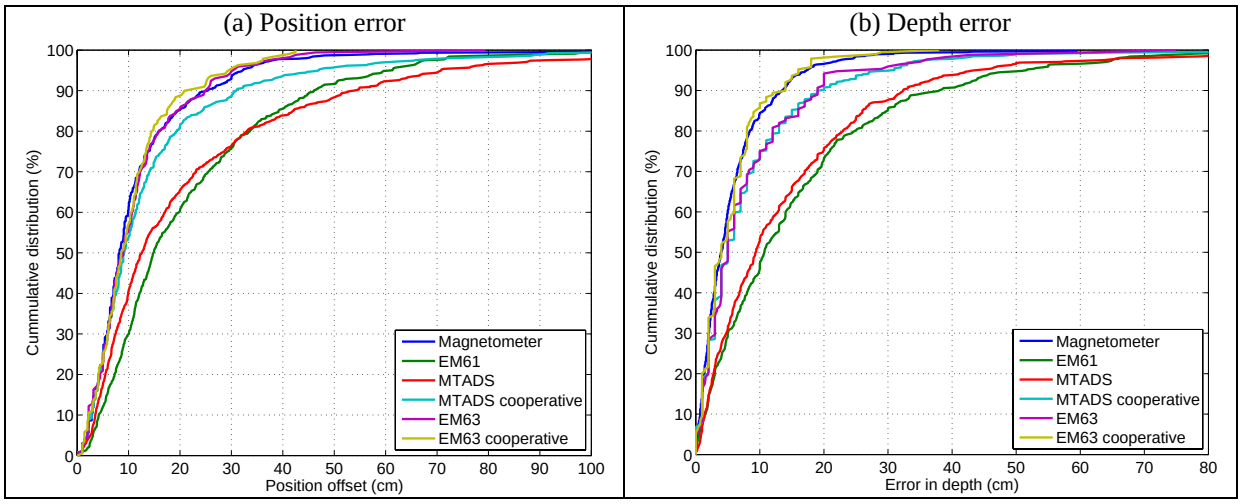



**Figure 17.** Cumulative distributions of (a) position; and (b) depth errors.

## 4.5 Comparison of single and cooperative inversion parameters

Our performance objective was to demonstrate that the within class variance of the cooperative inversion were less than that of the MTADS EM61 and EM63 data inverted without magnetometer constraints. Figures 17 and and b demonstrated that both the positions and depths estimates were improved markedly when the data were inverted cooperatively. Table 10 presents the standard deviations of the 4.2" mortar class (all fits, including the GPO are included). For both the MTADS EM61 and the EM63 there is a significant reduction in the within class variance of the size-based parameters. For the MTADS EM61 cooperative there is a slight increase of the within-class variance for the time-decay parameter, while for the EM63 there is a significant reduction.

**Table 10.** Standard deviations of the size and time-decay parameters for the 4.2" mortars.

| Sensor | Size parameter | Time-decay parameter |
|---|---|---|
| MTADS EM61 | 0.132 | 0.0134 |
| MTADS EM61 cooperative | 0.096 | 0.0145 |
| EM63 | 0.250 | 0.0188 |
| EM63 cooperative | 0.102 | 0.0129 |

## 4.6 Processing time

During the demonstration, each analyst kept a log of the time they spent on each step in the inversion process (Table 11). These steps included:

- Preprocessing and setup: This included all the time spent manipulating and importing into UXOLab all the data provided by each data collection demonstrator. Times ranged from 4 hours for the EM63 to 40 hours for the contractor EM61. A significant amount of time had to be spent on the EM61 cart data as the delivered data were not as well organized as the MTADS EM61 and magnetometer data. For the cooperative inversion we assumed the times would be the same as for single-inversion.

- Defining masks: This included the time spent automatically generating and then reviewing each mask prior to inversion. Typically this required 6 to 8.5 hours of analyst time for the full-coverage datasets. We assumed masks times for cooperative and single inversion were identical.
- Inversion: The computer time required to invert all anomalies which ranged from 3 hours for the magnetics to 15 hours for the MTADS EM61. The cooperatively inverted datasets required less computer time as many fewer start-models were required.
- PDF report generation: The computer time required to create PDF reports for each dataset. These are rough estimates as many of these were left to run overnight.
- QC of fits using the PDF report: The analyst reviewed each anomaly and determined if the inversion result could be trusted. Times ranged from 7 hours for the MTADS EM61 cooperative to 26 hours for the EM61. The better the data quality the less time required for QC.
- Mask and invert 2nd pass: Required between 6 to 8.5 hours for the full-coverage datasets. No remasking was required for the EM63 cooperative.
- QC using 2nd PDF report: Ranged from 4 to 10.5 hours for the full-coverage datasets.
- Independent QC of fits: The QC operator spent between 8 and 10 hours on quality control of the fitted data. This included any time required to remask and reinvert anomalies that failed the QC check.
- Discrimination: The selection of feature vectors, training of the classifier, application of the classifier and creation of the dig-sheet typically required about 6 hours per dataset.

**Table 11.** Time spent processing each of the different sensor combinations.

| Operation | Magnet-meter | EM61 | MTADS EM61 | MTADS EM61 coop. | EM63 | EM63 coop. |
|---|---|---|---|---|---|---|
| Preprocessing and setup | 11 | 40 | 18.25 | 18.25 | 4 | 4 |
| Defining Masks | 7 | 6.5 | 8.25 | 8.25 | 1.25 | 1.25 |
| Inversion | 3 | 6.5 | 15 | 4 | 7 | 2 |
| Generating PDF report | 20 | 6.5 | 18.5 | 18.5 | 2.25 | 2.25 |
| QC of fits using PDF report | 18.75 | 26 | 15.75 | 7 | 4 | 6 |
| Mask & invert 2nd pass | 7.5 | 8.5 | 6.5 | 6 | 0.5 | |
| Create 2nd PDF report | 5 | 5 | 3 | 4 | 1 | |
| QC using 2nd PDF report | 8 | 4.25 | 10.5 | 5 | 0.5 | |
| Independent QC of fits | 8.25 | 10 | 10 | 10 | 6 | 6 |
| Discrimination | 2 | 6 | 6 | 2 | 3 | 1 |
| **Total (hours)** | 90.5 | 119.3 | 111.8 | 83.0 | 29.5 | 22.5 |
| **Total minus computer time (hours)** | 62.5 | 101.3 | 75.3 | 56.5 | 19.3 | 18.3 |
| **Number of targets** | 1007 | 671 | 908 | 908 | 216 | 216 |
| **Per target (mins)** | 5.4 | 10.7 | 7.4 | 5.5 | 8.2 | 6.3 |
| **Per target minus computer time (mins)** | 3.7 | 9.1 | 5.0 | 3.7 | 5.3 | 5.1 |
| **Per target minus computer & setup (mins)** | 3.1 | 5.5 | 3.8 | 2.5 | 4.2 | 4.0 |

We had expected to be able to process and interpret each anomaly in less than 5 minutes. As shown in Table 11, if we include the set-up time, then this time goal was met with the magnetometer and MTADS EM61 data but not with any of the other datasets. The contractor EM61 required the most time with 9.1 minutes per anomaly. The MTADS EM61 and EM63 cooperative inversions both required about 8.75 minutes per anomaly (the times in Table 11 did not include the time required to invert the magnetometer data). If we discount the time required for set-up (not unreasonable if the data are delivered in a more usable form), then the time goal is met by all the single inversion methods except the EM61. The cooperative inversions the require about 7 minutes per anomaly.

## 4.7 Examine performance after elimination of the "can't analyze Data Analysis, Interpretation and Evaluation

In this section of the report we:

1) Conduct a failure analysis for the false-negative in the EM63 dig-list;
2) Analyze the number and type of anomalies recommended for excavation by each method;
3) Examine performance after elimination of the "can't analyze" category by using an expanded FOM;
4) Investigate the use of a pre-screener to reduce the number of geological alarms;

### 4.7.1 Failure analysis for the EM63 false negative

The EM63 data, when inverted without the magnetometer depth constraints, caused one false-negative on anomaly number 649. This was a 4.2" mortar buried at 40 cm which was predicted to be at 5 cm depth. The shallower solution depth caused the size of the item to be underestimated and resulted in the feature vector lying outside the main cluster of the 4.2" mortar class. Inspection of the observed data and model fit reveals a very poor fit to one of the transverse lines in the cued-interrogation data (Figure 18). This line of bad data should have been identified and removed so that the false-negative is a consequence of a quality control failure. A misfit versus depth plot (Figure 19) reveals that there are two minima: one at the shallow solution and another close to the true depth of 40 cm. The shallower solution has a slightly lower misfit and hence it is selected by the optimization algorithm. After removing the corrupt line of data, the misfit of the shallow solution increases, and the deeper solution is preferred (Figure 19). A better size estimate results and the feature vector for the anomaly lies within the same tight-cluster as the rest of the 4.2" mortars. When the depth constraints from the magnetometer were used, the deeper solution was also preferred and the false-negative was eliminated. This example demonstrates the importance of careful QC, and the benefits of the cooperative inversion procedure.

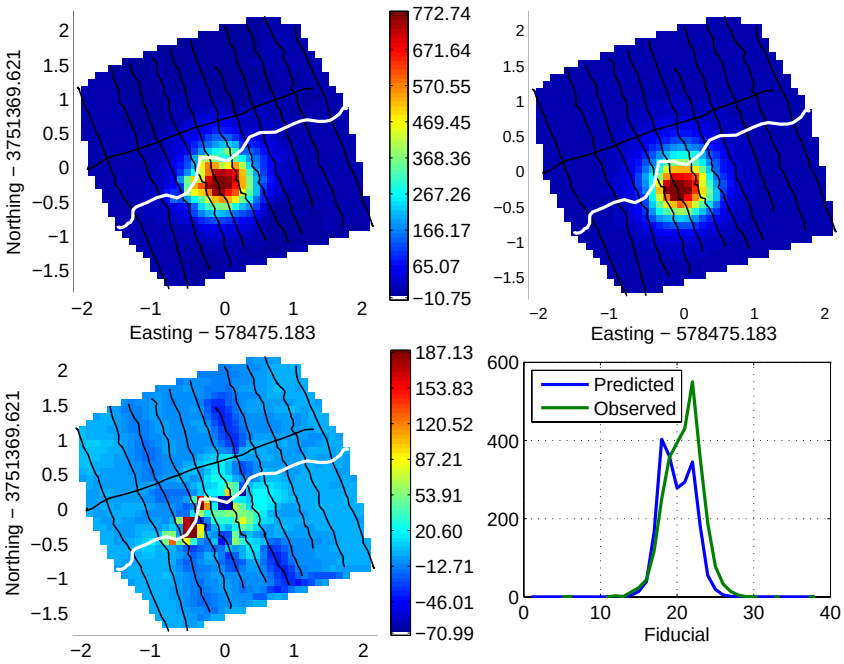



**Figure 18.** Data (top left), model (top right) and residual (bottom left) of the first time-channel dipole model fit to anomaly 649 in the EM63 data. On the bottom right is a profile of the actual and fitted data in time-channel 1 along the line shown in white.

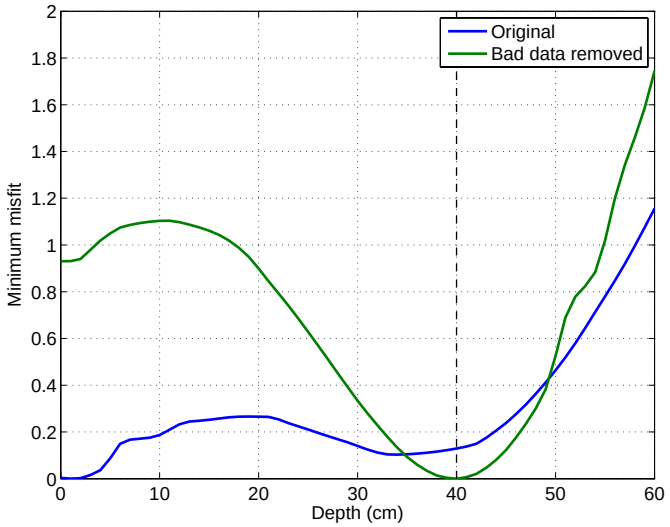



**Figure 19.** Misfit versus depth curve for the EM63 Pasion-Oldenburg model fit to anomaly 649. Two cases are considered: (i) using all the data which produces a shallow minimum close to the surface; (ii) after removal of a "bad" line of data, where the minimum occurs very close to the true depth of 40 cm (shown as a black dashed line)

### 4.7.2 Number and type of anomalies recommended for excavation

Figure 20 summarizes the percentage of each class of anomaly (junk, shrapnel, base-plates, partial rounds, UXO and geology) that need to be excavated at Pdisc = 1, at the classifier operating point and as “can’t analyze”. For the magnetometer, almost 90% of the partial rounds

were excavated by the time all UXO were recovered, compared to 25, 5 and 1% for the EM61, MTADS EM61 and MTADS EM61 cooperative respectively. These numbers demonstrate the superior intrinsic discrimination ability of the EM sensors compared to the magnetometer. Only a very small percentage of junk, base-plates and shrapnel were excavated at Pdisc = 1 for any of the methods. Between 35 to 75% of the partial rounds were excavated at the EM sensor operating points, indicating that the stop digging thresholds were too conservative. For the EM61 over half of the base-plates required excavation compared to 22 and 12% for the MTADS EM61 and MTADS EM61 cooperative. The better data quality of the MTADS EM61 produced more accurate feature vectors for the smaller sized items, and hence many did not need to be excavated as suspected UXO. The MTADS EM61 required excavation of less partial rounds but more base-plates than the MTADS EM61 cooperative.

Between 70-80% of the geological anomalies in the EM data were placed in the "can't analyze" category and had to be excavated. Larger percentages of the junk, shrapnel and base-plates were placed in the "can't analyze" categories for the EM61 and MTADS EM61 compared to the MTADS EM61 cooperatively inverted. The inclusion of the depth constraint from the magnetometer data removed some of the depth ambiguity and allowed the analyst to be less conservative in their pass/fail designations. Very few anomalies with a metallic source were listed as "can't analyze" in the magnetometer dig-sheet.

### 4.7.3 Elimination of "can't analyze" category using the Figure of Merit

For the MTADS EM61 and MTADS EM61 cooperative dig-lists a large number (285 and 205 respectively) of "can't analyze" anomalies had to be excavated. The "can't analyze" category was defined as follows. First, an automated analysis was conducted on a number of data and inversion features and each anomaly was assigned a Figure of Merit. FOM=0 when any one of the four metrics considered where less than a user defined threshold, and FOM=1 when all metrics were above the threshold. Second (and independent of the FOM analysis), an experienced interpreter conducted a visual review of all inversion results and provided an expert opinion as to whether the inversion result should be accepted or rejected. This opinion was based on a number of factors including whether the model accurately fit the data, the variation of misfit versus depth, the data coverage and any signal in the model outside the area of the mask. "Can't analyze" anomalies were those that the interpreter rejected AND which had FOM = 0. Inversion results that were passed by the interpreter and with FOM = 0 or inversion results that were failed with FOM = 1, were treated as "poor-fit" anomalies and were classified with a conservative threshold. Anomalies that were passed by the interpreter and with FOM = 1 were deemed "high-quality fits" and were classified with a more aggressive threshold.

As an alternative to the "can't analyze" category, we could have just used the FOM analysis as described above with one small modification: any anomaly that the analyst classified as fail would be set to FOM=0. Figure 21 shows the ROC curves that result, when the dig-list is ordered without a "can't analyze" category and with the two different classification thresholds based on the FOM. For all three datasets, Pdisc=1 at the discrimination operating point. For both MTADS EM61 datasets, all 4.2" mortars are classified as "high-confidence UXO" and the variable classification thresholds don't impact performance. One 4.2" mortar in the EM61 diglist has

FOM=0 and occurs between the low and high FOM classification cutoffs. In that case, the less aggressive cutoff for FOM=0 prevents a false-negative from occurring.

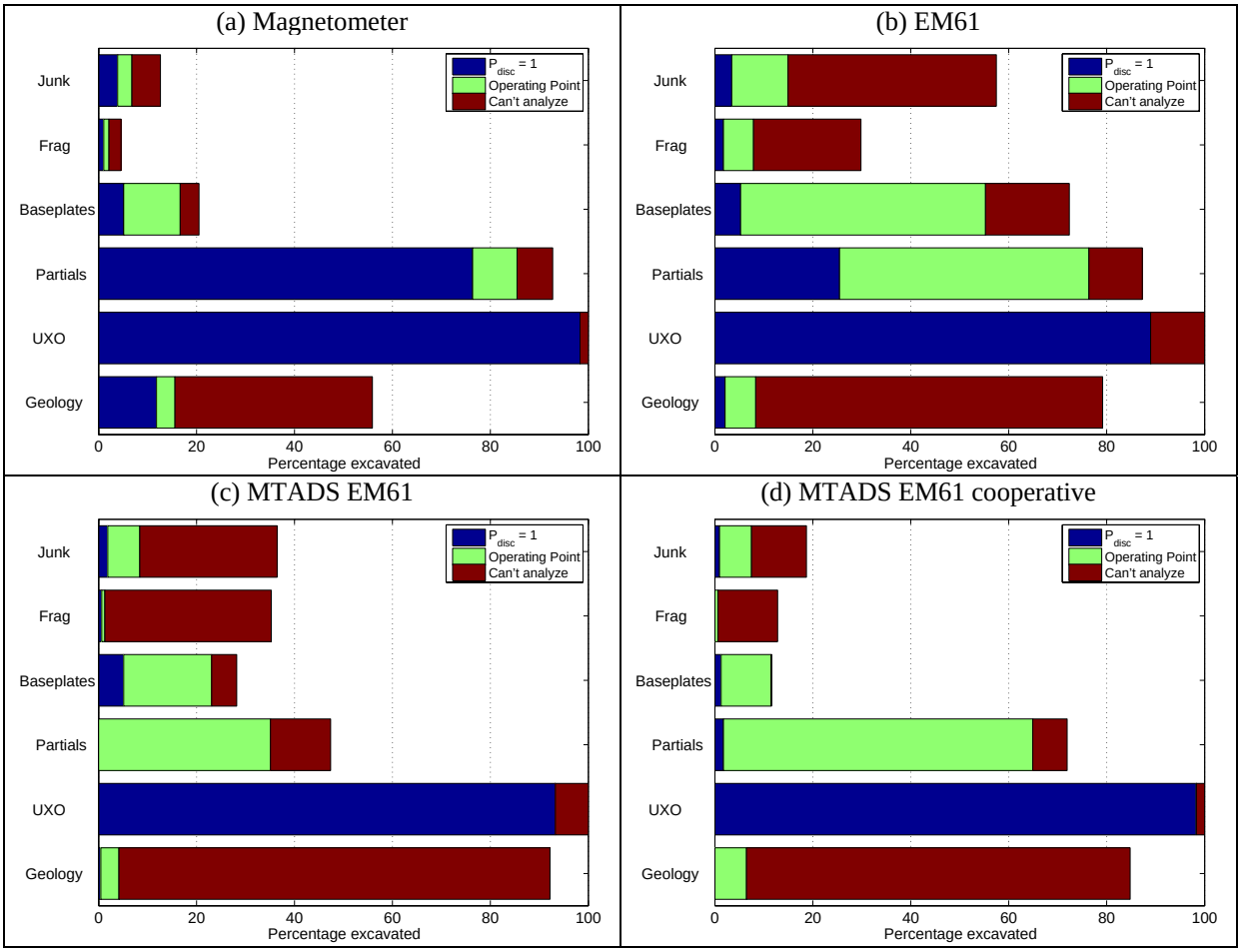



**Figure 20.** Comparison of the percentage of items of each class that were excavated at the classifier operating point, at the point where Pdisc = 1 and as "can't analyze".

Table 12 summarizes the number of false-positives for the EM61 cart, MTADS EM61 and MTADS EM61 cooperative when replacing the "can't analyze" category with the FOM adjusted classification boundaries. The FOM method would have reduced the number of excavations by 59, 161 and 101 anomalies for the EM61, MTADS EM61 and MTADS EM61 cooperative respectively. At the point where Pdisc = 1, there were 179, 65 and 32 false-positives, compared to 174, 293 and 240, respectively. This means there was a slight increase in the number of FP for the EM61, and significant reductions in the numbers for the MTADS EM61 and MTADS EM61 cooperative.

**Table 12.** Comparison of the number of excavations required for the EM61, MTADS EM61 and MTADS EM61 cooperative when using FOM analysis in place of the "can't analyze category"

| Sensor | FP at original OP | Original # "can't analyze" | FP at OP using FOM | Reduction in FP | % decrease "can't analyze" | FP at Pdisc = 1 |
|---|---|---|---|---|---|---|
| EM61 | 264 | 149 | 205 | 59 | 40% | 179 |
| MTADS EM61 | 344 | 285 | 183 | 161 | 56% | 65 |
| MTADS EM61 cooperative | 275 | 205 | 174 | 101 | 49% | 32 |

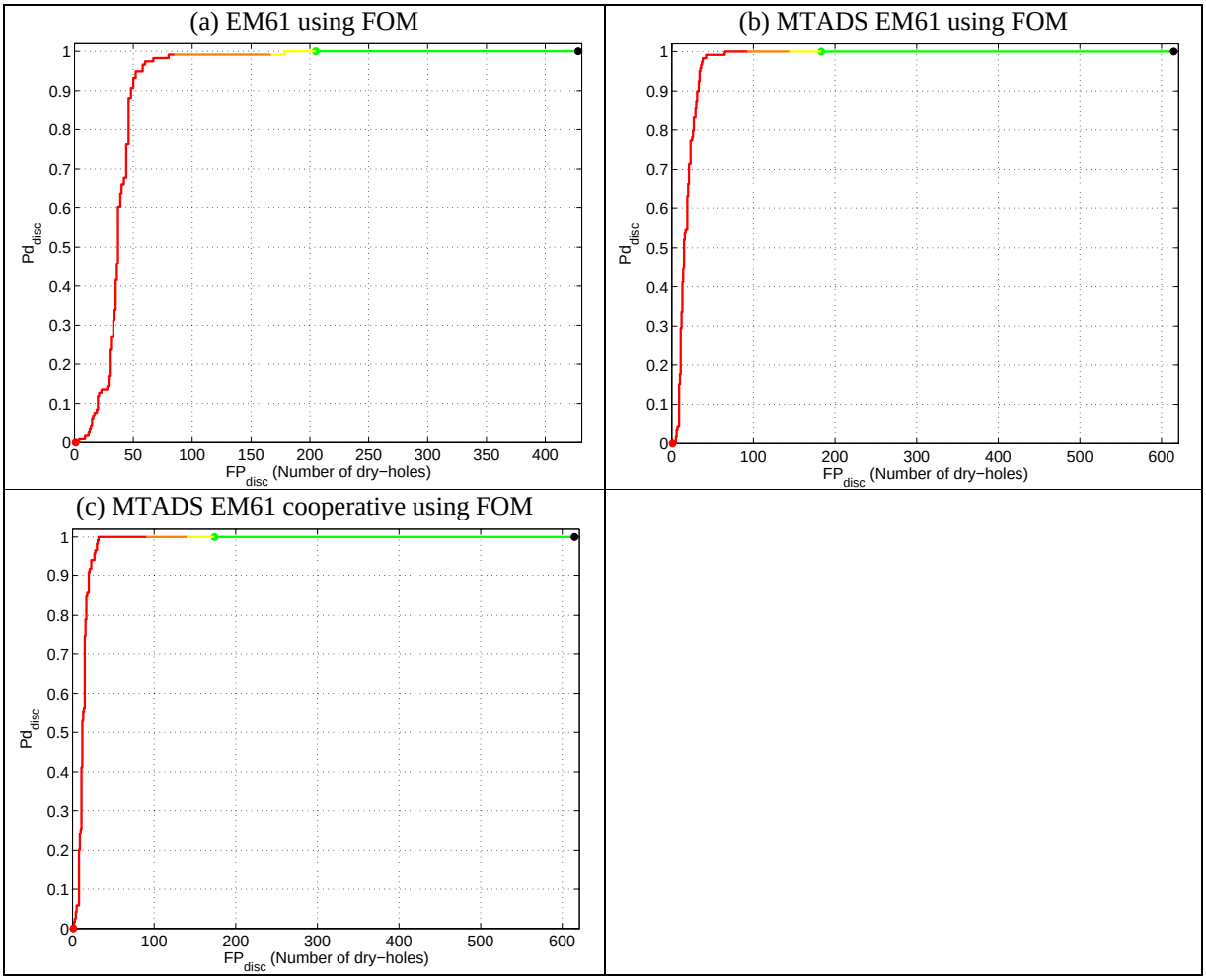



**Figure 21.** ROC curves for the EM61, MTADS EM61 and MTADS EM61 cooperative when the FOM is used in place of the "can't analyze" category.

### 4.7.4 Reduction of the number of geological false-alarms by a prescreener

A large number of the "geological" alarms detected by the MTADS EM61 array were due to bouncing of the cart as it traversed across furrows in the ploughed field on the South-West section of the site. This is apparent in a scatter-plot of the total energy of time-channel t1 of the East-West versus North-South transects (Figure 22a). Most of the geological anomalies have between 18-23 dB energy North-South compared to 5 to 18 dB East-West. To reduce the number of geological anomalies we decided to try to develop a "metal-soil" discriminator based on the relative energy in the two directions. This would be applied as a prescreener before submitting the anomaly to the polarization tensor fitting routines. We used the GPO and initial training data to investigate different potential classifiers and used two classes: one comprising the geological anomalies, and the other all the metallic anomalies minus the fragmentation. After some initial experimentation we settled on a quadratic discriminant analysis classifier as it produced the most intuitively reasonable discrimination boundary. We then plotted the cumulative distributions of the different classes against the "metal" probability, $P_{metal}$ (Figure 22c). Approximately 50% of geological anomalies had $P_{metal} < 0.3$ with the first UXO occurring at $P_{metal} = 0.4$. We therefore selected an operating point of $P_{metal} < 0.3$.

Figure 22b shows a feature vector plot of the blind-test data and Figure 22d shows cumulative distributions of the different classes. At the selected operating point, 55% of geological anomalies would be rejected with 14% of shrapnel, 8% of scrap metal and less than 3% of partial rounds and base-plates excluded from further analysis. Most importantly, no UXO would be rejected by the pre-screener. It would have allowed 119 geological anomalies to be excluded from further analysis: 116 of these were in the "can't analyze" category and had to be excavated. In total, the pre-screener could have reduced the number of "can't analyze" anomalies by 130 (from 285 down to 155). For the MTADS EM61 cooperative inversion the reduction would have been from 226 down to 116.

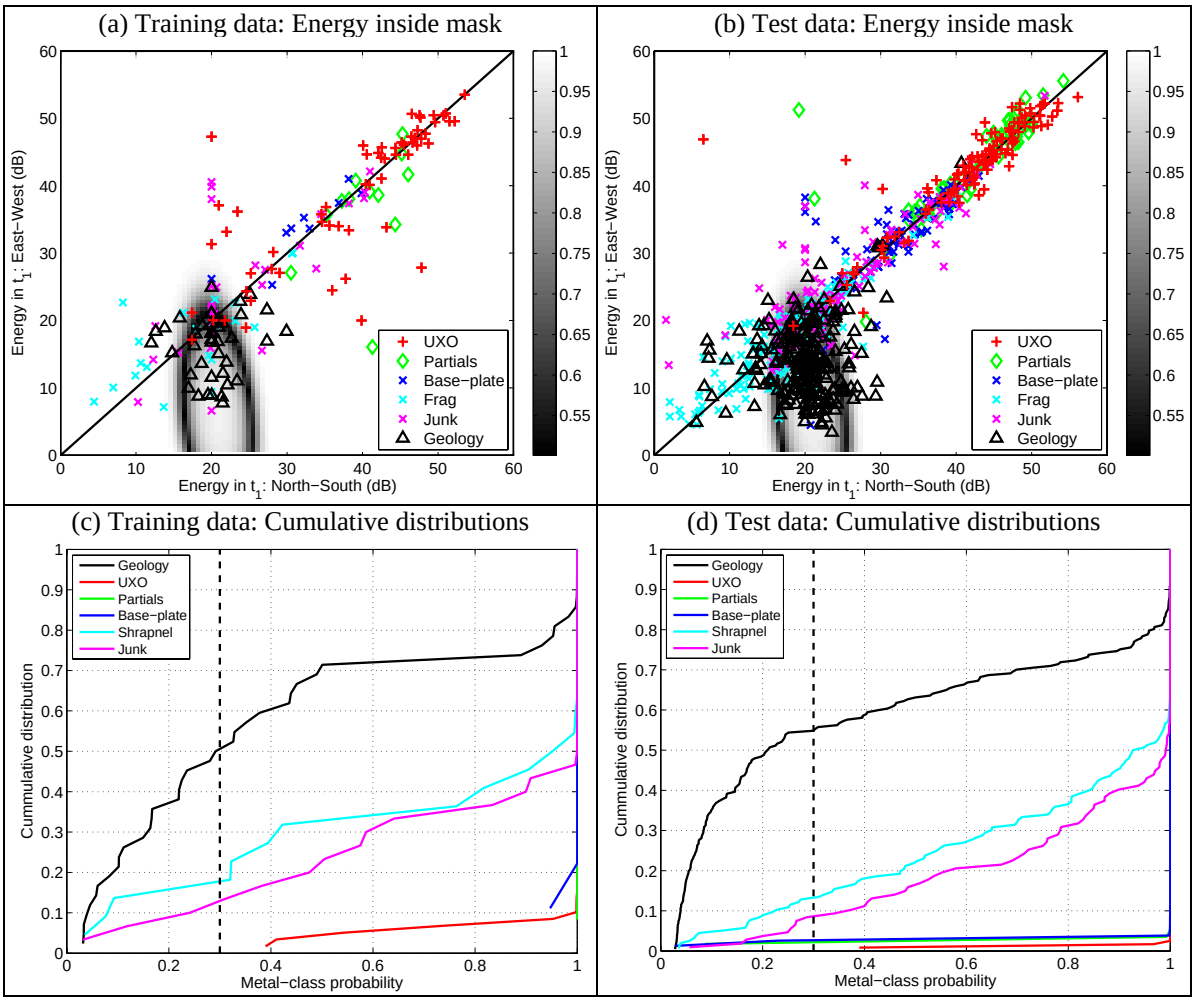



**Figure 22.** Reduction in geological false-alarms in the MTADS EM61 data using the difference in the energy in the North-South versus East-West lines: Energy feature vectors from the (a) training and (b) test-data overlying the classifier decision surface; and cumulative distributions of classes on the (c) training and (d) test-data when ranked by "metal" probability.

# 5 COST ASSESSMENT

## 5.1 Cost Reporting

The demonstration costs for each of the different methods were tracked throughout the demonstration (Table 13). Only approximate costs and times were available for the one UBC employee (LinPing Song) who participated in the data analysis component of the project. The magnetometer interpretation required less time and cost than the other full-coverage methods. The MTADS EM61 required less time to interpret than the contractor EM61, partly because the EM61 data required a more time to organize and manage (due to the way it was delivered to demonstration participants). For cooperative inversion we list two costs. The first represents the additional costs associated with inversion, QC and interpretation using the magnetometer locations as constraints. However, these are not a good indication of the true cost of cooperative inversion as the anomalies had already been preprocessed and masked prior to inversion. To estimate the cost of cooperative inversion we add the cost of the magnetometer interpretation (calculated as the number of EM anomalies times the cost per anomaly for magnetometer interpretation) to the cost of the single inversion interpretation of the EM data. These numbers are presented in parentheses.

**Table 13.** Cost Summary. For the cooperative inversion dig-sheets the first number represents the additional cost of cooperative inversion (after single inversions have been complete). The second number in parentheses represents the estimated total cost of cooperative inversion (including the magnetometer interpretation).

| Category | Hours | Cost | # anomalies | Cost per anomaly |
|---|---|---|---|---|
| Pre-Demo Testing Prep | 542 | *$62,652* | | |
| MTADS EM61 Interpretation | 165 | *$15,259* | *870* | *$17.54* |
| Magnetometer Interpretation | 82 | *$8,167* | *969* | *$8.43* |
| EM63 Interpretation | 65 | *$6,431* | *178* | *$36.13* |
| Contractor EM61 Interpretation | 186 | *$18,685* | *633* | *$29.52* |
| MTADS EM61 Cooperative | 46 | *$4,966 ($22,593)* | *870* | *$5.71 ($25.97)* |
| EM63 Cooperative | 16 | *$1,491 ($6755)* | *178* | *$8.38 ($37.95)* |
| Demonstration Report | 228 | *$38,646* | | |
| **Total** | **1339** | **$156,297** | | |

## 5.2 Cost Analysis

### 5.2.1 Cost Comparison

We anticipate that the Program Office will conduct a cost comparison of the different discrimination methods.

### 5.2.2 Cost Basis

The anticipated cost basis for this technology demonstration is the number of anomalies per day which can be processed and interpreted.

### 5.2.3 Cost Drivers

The main factor expected to be a strong cost driver for this technology is the number of anomalies that can be processed per day.

# 6 REFERENCES

Bell, T. H., B. J. Barrow, and J. T. Miller, 2001, Subsurface discrimination using electromagnetic induction sensor, *IEEE Transactions on Geoscience and Remote Sensing*, 39, 1286 – 1293.

Billings, S. D., L. R. Pasion, and D. W. Oldenburg, 2002b, Discrimination and identification of UXO by geophysical inversion. Phase II: Inversion of total-field magnetics, *USACE Engineer Research and Development Center*, Tech. Rep. ERDC/GSL TR-02-16.

Billings, S.D., 2004, Discrimination and classification of buried unexploded ordnance using magnetometry: *IEEE Transactions of Geoscience and Remote Sensing*, 42, 1241 - 1251.

Collins, L., Zhang, Y., Li, J., Wang, H., Carin, L., Hart, S., Rose-Phersson, S., Nelson H. & McDonald, J. R., 2001, A comparison of the performance of statistical and fuzzy algorithms for unexploded ordnance detection: *IEEE Transactions on Fuzzy Systems*, 9, 17-30.

Carin, L., Zhang, Y. & Liao, X., 2004, Detection of Buried UXO via Active Selection of Labeled Data, proceedings from the UXO Forum, St. Louis, March 9-12, 2004.

ESTCP, 2006, ESTCP Discrimination Study Demonstration Plan, Draft 2, October

Hart, S. J. et al., 2001, Using Physics Based Modeler Outputs to Train Probabilistic Neural Networks for Unexploded Ordnance (UXO) Classification in Magnetometry Surveys. *IEEE Trans. Geosci. Remote Sensing* 39, 797-804.

Pasion, L. & Oldenburg, D. 2001, A Discrimination Algorithm for UXO Using Time Domain electromagnetics: *Journal of Engineering and Environmental Geophysics*, 28, 91-102.

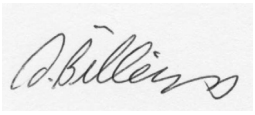
Pasion, L. R., Billings, S. D. and Oldenburg, D. W., 2003, Joint and Cooperative Inversion of Magnetics and Electromagnetic Data for the Characterization of UXO Discrimination Problems: *The Symposium on the Application Geophysics to Engineering and Environmental Problems (SAGEEP),* San Antonio.

Zhang, Y., Collins, L. M., & Carin, L., 2003a, Model-based statistical signal processing for UXO discrimination: Performance results from the JPG-V demonstration: *Proceedings of SPIE Volume 5089,* 1116-1126.

Zhang, Y., Collins, L., Yu, H., Baum, C. E. & Carin, L., 2003b, Sensing of Unexploded Ordnance with Magnetometer and Induction Data: Theory and Signal Processing. *IEEE Trans. Geosci. Remote Sensing,* 41, 1005-1015.

# 7 POINTS OF CONTACT

| | | |
|---|---|---|
| **ESTCP** | | |
| Anne Andrews | Program Manager, MM | Tel: 703-696-3826<br>Fax: 703-696-2114<br>anne.andrews@osd.mil |
| Katherine Kaye | Program Assistant, MM | Tel: 410-884-4447<br>Fax: 703-478-0526<br>kkaye@hgl.com |
| **Sky Research, Inc.** | | |
| Stephen Billings | Principal Investigator | Tel: 541-552-5185<br>Cell: 604-506-9206<br>stephen.billings@skyresearch.com |
| Joy Rogalla | Project Manager | Tel: 541-552-5104<br>Cell: 541-292-1653<br>joy.rogalla@skyresearch.com |



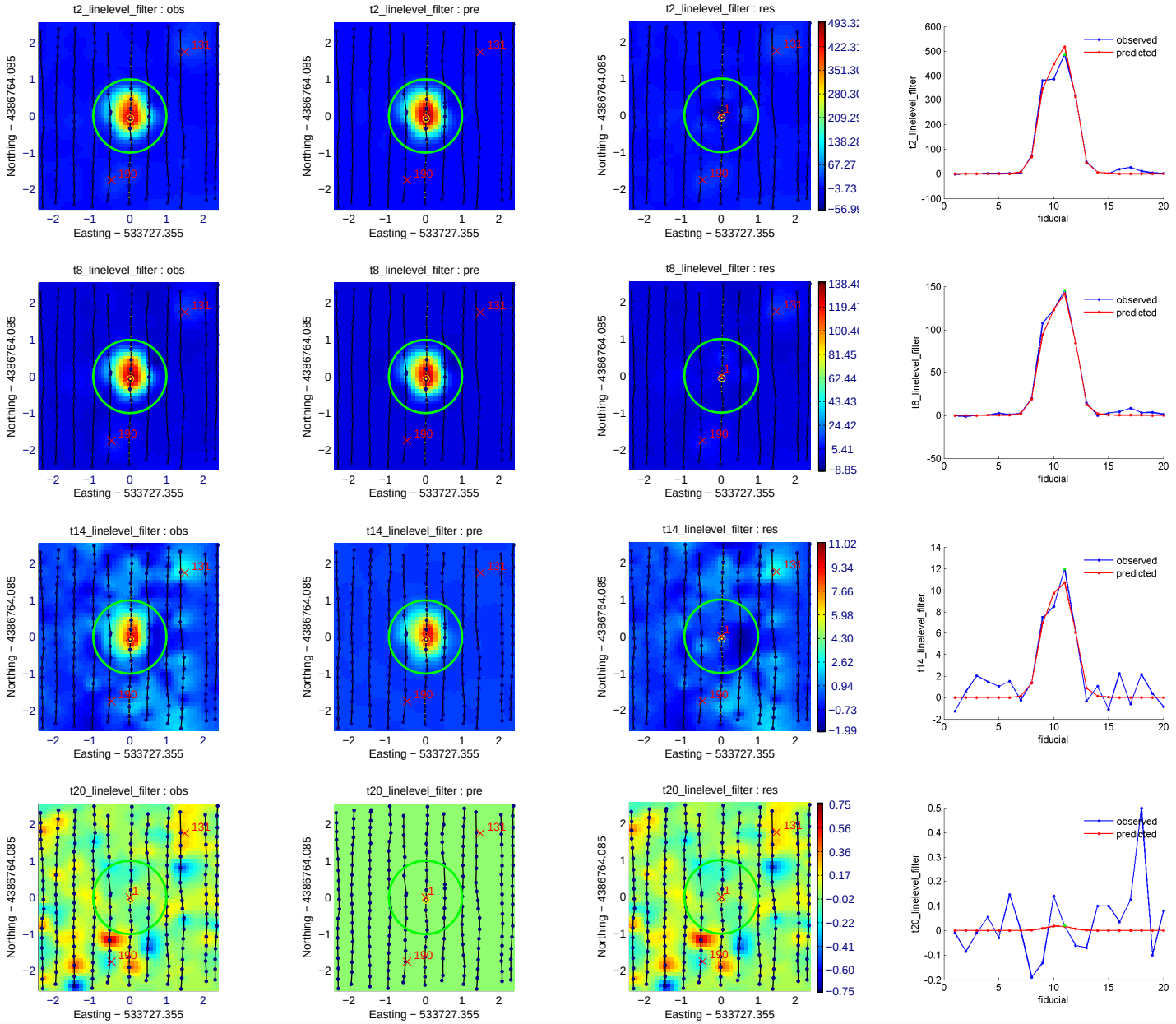
**Stephen Billings**
**September 5, 2008**

# APPENDIX A

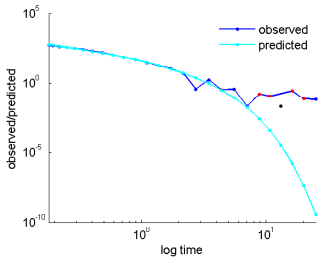
## Examples of PDF Inversion Reports

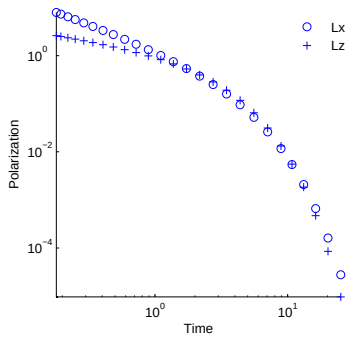
# 1 Cell 1

Fit? None Save Fits Contents Next Cell




| | $L_x$ | $L_z$ |
|---|---|---|
| L 0.195 | 7.271 | 27.2836 |
| L 0.22 | 6.1997 | 23.2502 |
| L 0.25 | 5.22 | 19.5559 |
| L 0.29 | 4.2557 | 15.9141 |
| L 0.34 | 3.3976 | 12.6698 |
| L 0.405 | 2.629 | 9.7618 |
| L 0.485 | 1.9955 | 7.3653 |
| L 0.59 | 1.4554 | 5.3254 |
| L 0.72 | 1.0347 | 3.7422 |
| L 0.89 | 0.69981 | 2.4906 |
| L 1.105 | 0.45267 | 1.5768 |
| L 1.375 | 0.27827 | 0.94251 |
| L 1.725 | 0.15796 | 0.51533 |
| L 2.17 | 0.082215 | 0.25544 |
| L 2.74 | 0.038173 | 0.11128 |
| L 3.465 | 0.015437 | 0.041445 |
| L 4.385 | 0.0052543 | 0.012691 |
| L 5.565 | 0.0014181 | 0.0029869 |
| L 7.07 | 0.00028718 | 0.00050733 |
| L 8.905 | 4.3931e-005 | 6.255e-005 |
| L 10.78 | 6.7877e-006 | 7.7459e-006 |
| L 13.22 | 6.2857e-007 | 5.3733e-007 |
| L 16.28 | 3.3573e-008 | 1.9958e-008 |
| L 20.15 | 8.7413e-010 | 3.2792e-010 |
| L 25.14 | 8.4181e-012 | 1.7423e-012 |



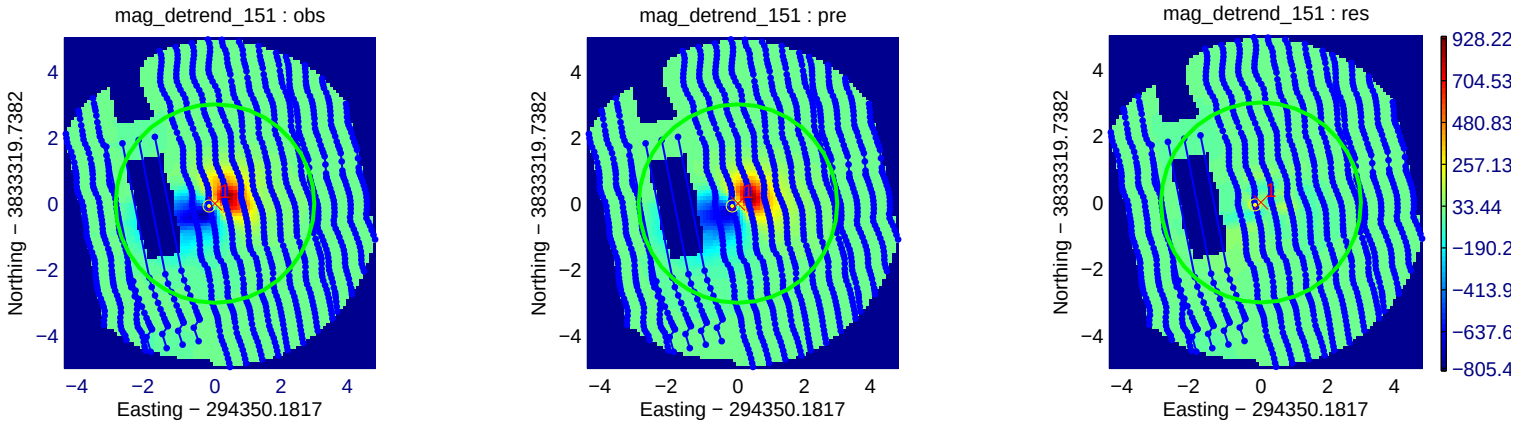



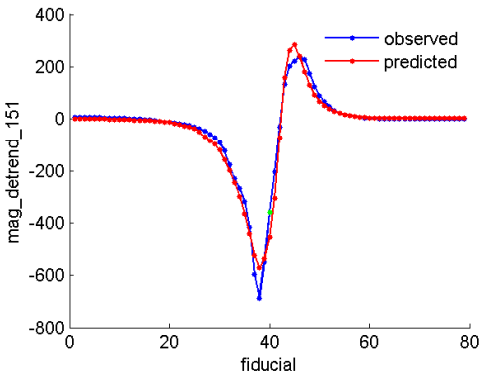

| Easting | Northing | Depth | $\theta$ | $\phi$ | $\kappa_1$ | $\beta_1$ | $\gamma_1$ | $\kappa_2$ | $\beta_2$ | $\gamma_2$ | misfit | CorrCoef |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 533727.2954 | 4386764.1663 | -0.017732 | 55.0006 | 91.6804 | 5.3234 | 1.1189 | 1 | 1.3411 | 1.139 | 1.1365 | 1.4824 | 0.84233 |

| firstorderopt | iterations | funcCount | cgiterations | algorithm | message | numstartmodels | time | S2N |
|---|---|---|---|---|---|---|---|---|
| 55.9625 | 14 | 180 | 59 | large-scale: trust-region reflective Newton | Optimization terminated: relative function value changing by less than OPTIONS.TolFun. | 10 | 20.1892 | 20 of 26 channels/freqs inverted. Minimum S2N = 1.00 |

# 1 Cell 1

Fit? None Save Fits Contents Next Cell




| Easting | Northing | Depth | m | $\phi$ | $\theta$ | dc | misfit | CorrCoef |
|---|---|---|---|---|---|---|---|---|
| 294350.1256 | 3833319.7974 | -0.29046 | 5.4222 | -124.6053 | 1.0166 | 2.8305 | 223236.0632 | 0.99103 |

| firstorderopt | iterations | funcCount | cgiterations | algorithm | message |
|---|---|---|---|---|---|
| 2450.5955 | 120 | 121 | 359 | large-scale: trust-region reflective Newton | Optimization terminated: relative function value changing by less than OPTIONS.TolFun. |

# APPENDIX B

## White Paper: Data Modeling, Feature Extraction, and Classification of Magnetic and EMI Data, ESTCP Discrimination Study, Camp Sibert, AL

## Project 200504: Practical Discrimination Strategies for Application to Live Sites

# INTRODUCTION

As part of the ESTCP UXO Discrimination Study, Sky Research and UBC-GIF will submit the following 8 dig-sheets:

a) Magnetics, size-based: Calculation of the magnetic remanence metric and the production of a dig-sheet ranked according to remanence;

b) MTADS EM-61, size-based: Production of a dig-sheet ranked according to size (using the sum of the beta parameters for time-channel 1 from the MTADS EM-61 data).

c) Contractor EM-61, size-based: Production of a dig-sheet ranked according to size (using the sum of the beta parameters for time-channel 1 from the Contractor EM-61 data).

d) MTADS EM-61, statistical: Statistical classification of features derived from the MTADS EM-61 data and the production of a ranked dig-sheet;

e) Contractor EM-61, statistical: Statistical classification of features derived from the Contractor EM-61 data and the production of a ranked dig-sheet;

f) EM-63, statistical: Same as b) but with the EM-63;

g) MTADS EM-61 and magnetics, statistical: As per b) but with EM-61 fits constrained by the magnetics data and with the addition of the features from the magnetometer data (remanence, moment etc); and

h) EM-63 and magnetics, statistical: As per d) but with the EM-63.

This document describes the fitting parameters used for each data type, and discusses the ranking strategy for each method.

# 1. MTADS MAGNETOMETER DATA

## a. Analysis of Geopyhsical Prove Out results

The following parameters were used to invert the magnetometer data from the Geophysical Prove Out (GPO):

- Earth's magenetic field: Inclination = 63.48$^o$, Declination = -3.40 $^o$, Magnitude = 50,700 nT;
- Noise-floor = 1.7 nT (from statistics on northern part of GPO) plus 2% of the magnitude of the signal;
- Elliptical mask based on a derived total-gradient channel (often referred to as analytic signal) with starting contour at 2.5% of the maximum value.

All dipole model fits were found to be acceptable except for targets 10 and 14 (1.1 and 1.18 m depths, respectively). The recovered moments in directions parallel and perpendicular to the Earth's magnetic field are plotted in Figure 1a, along with a first guess at the "dipole feasibility curve" for a 4.2" mortar (a spheroid of diameter 4.2" = 107 mm, and aspect ratio 3.73). All except four of the moments cluster around the dipole feasibility curve. A better estimate of the dipole feasibility curve can be obtained by solving for an equivalent spheroid that minimizes the discrepancy between the inverted and predicted dipole moments (Figure 1b). The predicted moments were obtained from the burial depths and orientations of each item on the GPO (neglecting the four items with significant remanence).

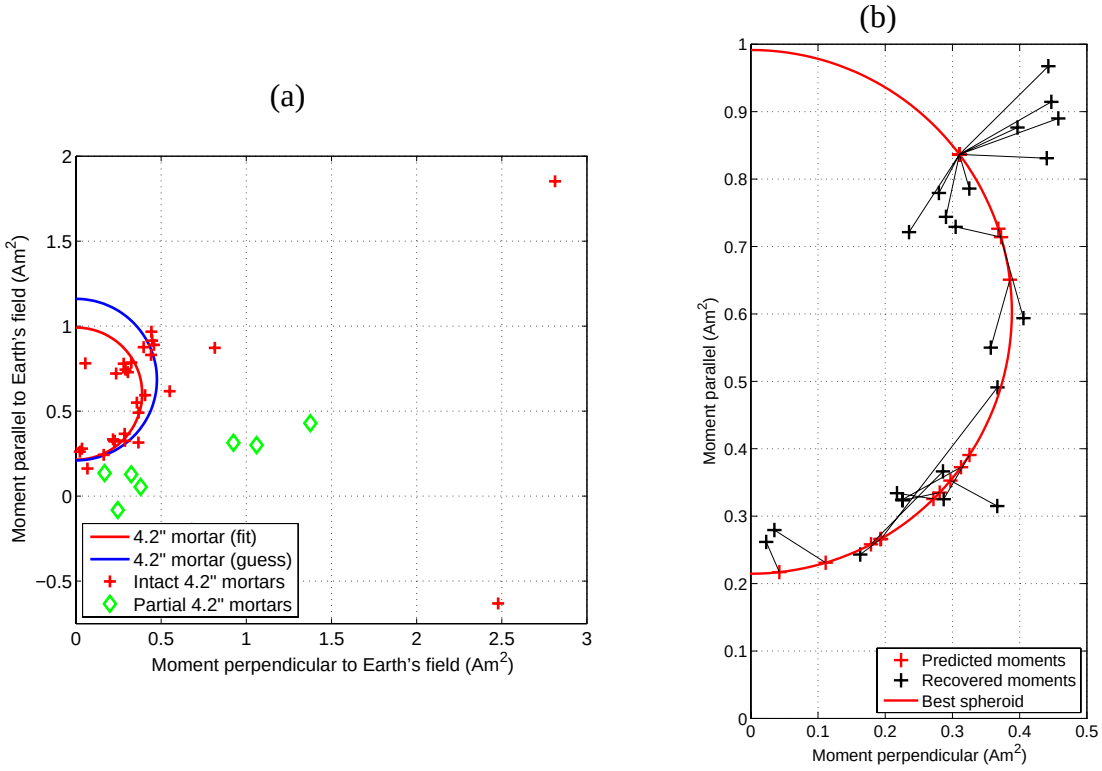



**Figure 1: (a)** Fitted magnetic moments from the GPO. **(b)** Dipole feasibility curve from the best-fitting spheroid, which shows the predicted and actual moments for each item (joined by a solid black line) in the GPO (neglecting the four outliers).

Two of the intact GPO items have moderate remanence (> 30%) and two have significant remanence (> 50%) with a maximum value of 80%. All other intact GPO items have relatively low remanence, with thirteen less than 10% remanance and eleven between 10-30% remanence. The partial rounds have large contributions to the moment that are perpendicular to the Earth's magnetic field, indicating a substantial remanent magnetization.

There is excellent agreement between predicted and actual depths and locations (Figure 2). All items are located within 20 cm of the ground-truth location, with a bias of 1 cm to the West and 6 cm to the South. All predicted depths are within 26 cm, and all but two within 10 cm of the actual depths. There is a bias of 4.3 cm resulting in a tendency to predict slightly deeper depths than the reported groundtruth.

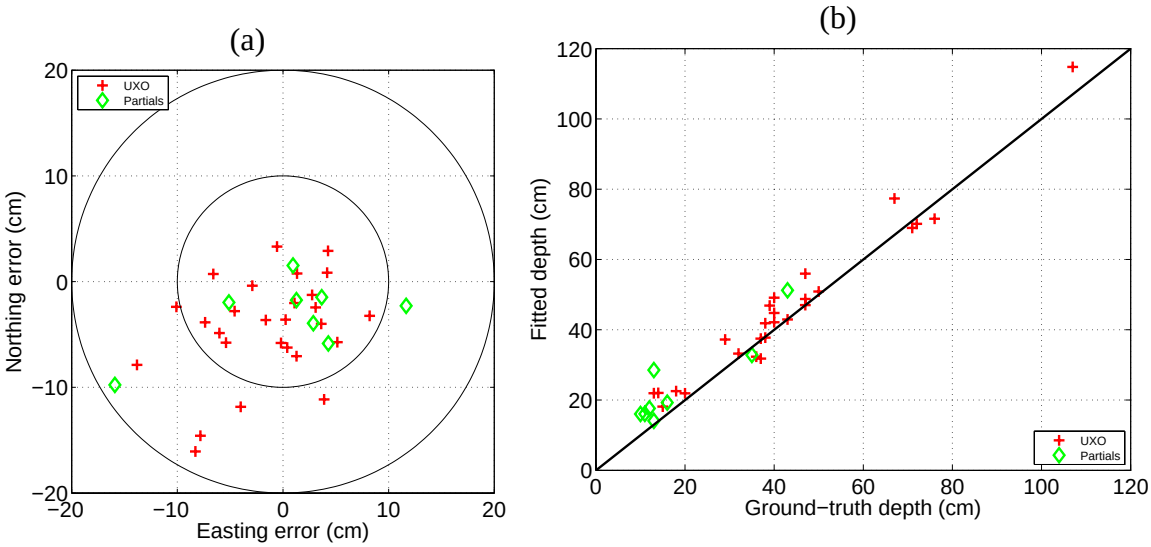



**Figure 2: (a)** Error in estimated dipole locations; **(b)** Predicted versus actual depths.

### b. Analysis of ground-truth data

The rest of the results were inverted using the same procedures as the GPO. There were 969 inverted anomalies, with 861 having acceptable dipole fit and 108 having an unacceptable fit. The ground-truth for 144 of the magnetometer anomalies was provided by the Program Office in early August, with 133 of the anomalies having valid dipole fits. Twenty-six of the ground-truth items were from seeded 4.2" mortars (there were two additional deep rounds in the ground-truth that were not detected). The dipole parameters and ground-truth information are summarized in Figure 3. Figures 3a and b show that all but four of the 4.2" mortars cluster around the dipole feasibility curve calculated from the GPO. One of the outliers has such a large remanence that its magnetization is reversed relative to the direction of the Earth's magnetic field. A scatter plot (Figure 3c) showing remanence versus moment indicates that either parameter would provide a good basis for a discrimination strategy. This is reflected in the ROC curve (Figure 3d), where all UXO's are recovered by both methods with small false-alarm rates (7 non-UXO for moment and 9 for remanence).

Figures 3e and f show that the recovered locations and depths agree very well with the reported ground-truth locations. Note that there is a tendency to predict shallower depths for the cluster

of seed items at 60 cm. We suspect the items were actually buried a little shallower as all of the EM methods have a similar bias.

### c. Analysis of ground-truth data

With the potential for significant remanence in the SEED items, we will opt to use the moment to prioritize digging order. We take the viewpoint that if the moment is small that the item cannot possibly be an intact 4.2" mortar. We will also provide an extra column in the digsheet which expresses the ranking order if we had used remanence so a retrospective performance analysis can be conducted. For the GPO items, the lowest recovered moment for a 4.2" round was 0.175 $Am^2$, while for the ground-truth it was 0.185 $Am^2$. Thus for our ranking we use a value of 0.17 $Am^2$ as our "high probability of UXO cutoff". To set our thresholds for "unable to make a decision dig" for most likely UXO and most likely not UXO, we use depth recovery errors of 2.5 and 5 cm. For the worst case of a shallow item pushed to the surface, this suggests conservative cutoff values of 0.14 and 0.109 $Am^2$, which we round down to 0.1 $Am^2$ to add an extra safety margin. That is, we will recommend digging all items with moments greater than 0.10 $Am^2$, but don't expect to find any UXO with moments less than 0.14 $Am^2$. We place the items that could not be fit between the largest moment < 0.14 $Am^2$ and the next largest item.

It's difficult to provide a quantitative estimate of the interpretation confidence for the size-based analysis as it relies on expert judgment to set the digging thresholds. If the training dataset is representative[2] of the other items encountered at the site, we have a high (~99%) confidence that hazardous items will not be left in the ground. The confidence would be lowered slightly if the training dataset is not representative.

For remanence, the maximum value on the GPO was 93%, and within the training data it was 96%. From past experience at other sites, we have not encountered live-site ordnance (as opposed with seeded ordnance) with remanence > 70%. Therefore, we use that number as our high-confidence ordnance threshold. We then set "unable to make a decision" thresholds at 100% and 110% remanence. These were arbitrarily selected so that all seeds thus far encountered would be excavated, with the second threshold to include an extra safety margin.

In Figure 4, we plot moment versus remanence for the unlabelled items and the ground-truth UXO (not the GPO items), and delineate the high-confidence ordnance and digging thresholds.

[2] Representative in terms of SNR, amount of remanence, depth of burial, geology etc

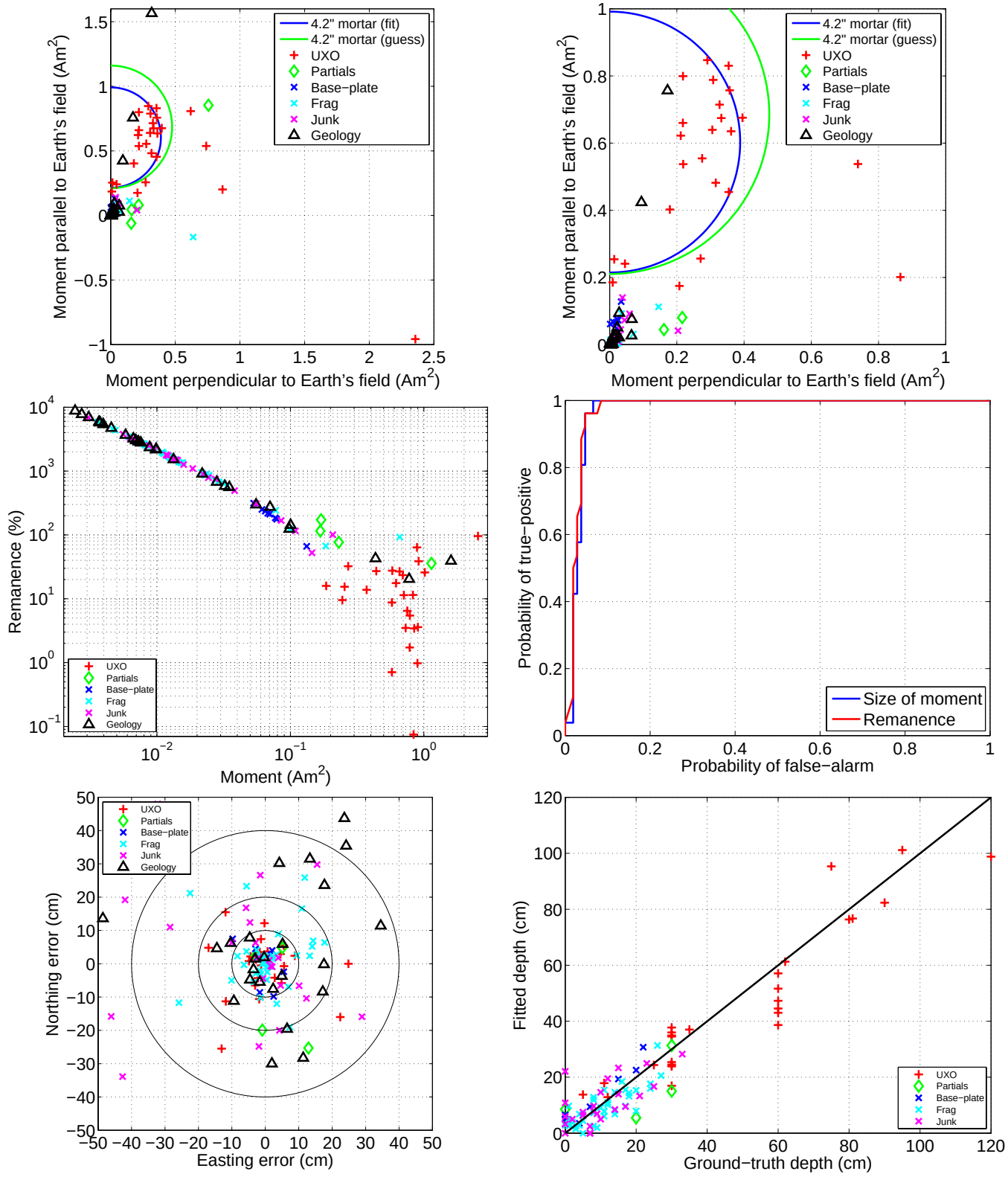



**Figure 3:** Dipole model fits to ground-truth items at Camp Sibert. From top-left to right, then down the plots are **(a)** moments perpendicular and parallel to Earth's field; **(b)** same as (a) but with reduced range; **(c)** scatterplot of moment versus remanence; **(d)** ROC curves based on the size of moment and the remanence; **(e)** error in the estimated dipole location; and **(f)** actual versus predicted depths.

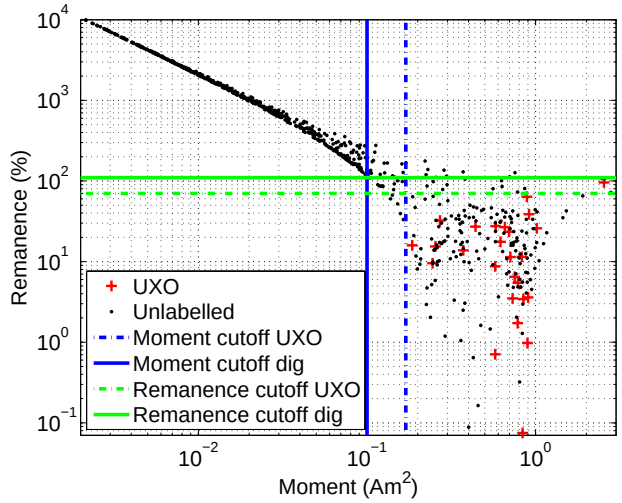



**Figure 4:** Moments and remanence of seeded UXO and unlabelled items, along with cutoff values for digging prioritized by moment and by remanence.

# 2. EM-61 CART DATA

## a. Analysis of Geophysical Prove Out results

Data were collected over thirty 4.2" mortars and eight partial mortars buried in the GPO with a Geonics EM-61 cart recording at 3 time gates on the bottom receiver plus a fourth channel made of the last time gate on the top receiver. We assumed that the geometry of the sensor was standard and set the sensor elevation to 0.5 m above the ground after a least-squares error analysis of predicted versus actual depth of buried items.

The demedian filtered data reported by Parson's were used for the inversions (that is, Sky Research did not perform any preprocessing of the data). The data were processed within UXOLab using the following workflow:

- Estimates of data errors consisted of two components:
  - Background noise was estimated by choosing data surrounding the anomaly, and calculating statistics for those data.
  - For each datum, we assumed an error equal to 10 percent noise of the amplitude.
- *Masking*: The spatial extent of data used in the inversions was determined using the elliptical masking technique described in the Demonstration Plan.
- *Model*: The data were fit using 3 unique polarizations for the dipole tensor. The amplitude of each polarization was estimated at each of the three time channels. The 15 element model vector is

$$m = \left[X, Y, Z, \phi, \theta, \psi, L_1(t_1), L_1(t_2), L_1(t_3), L_2(t_1), L_2(t_2), L_2(t_3), L_3(t_1), L_3(t_2), L_3(t_3)\right]$$

  where $(X,Y,Z)$ is the location, $(\phi,\theta,\psi)$ are the orientation angles and $L_i(t_j)$ is the $i$-th polarization at the jth time channel.

Note that the top-coil data (corresponding to the third-time channel) were used within the inversions so that there were four "channels" of data to constrain the instantaneous polarizations at the three time-channels. This extra information improved the depth estimates.

After inversion and quality control analysis of the GPO data, we obtain 28 successful inversions and 10 failed inversions. Inversions were disqualified when fits were deemed to be too poor to warrant advanced discrimination. Four of the five deep items failed for this reason.

Figure 5 plots different combinations of the recovered polarization parameters. Figure 5 (a) plots the primary polarization at the first time channel (i.e. $L_1(t_1)$) and the secondary polarizations at the first time channel (i.e. $L_2(t_1)$ and $L_3(t_1)$). This parameter plot gives an indication of the size of the target (since the polarization magnitude is proportional to size) and also the symmetry of the target. A target with axial symmetry would have equal secondary polarizations, i.e. $L_2(t) = L_3(t)$. The length of the y-directed error bar indicates the difference in $L_2(t)$ and $L_3(t)$. The data are unable to constrain the secondary polarizations to be equal for the 4.2" mortars indicating that any feature vectors so derived would be unreliable. The amplitude of the primary and secondary instantaneous polarizations at time-channel 1 provides reasonable separation between intact and partial 4.2" mortars, although there is some overlap between classes. The relative decay is defined to be the ratio of the polarizations at the 3rd and 1st time channels. Figure 5 (b) shows

that this parameter tends to be larger for the 4.2" mortars but again there is some overlap with the partial rounds. Figures 5 (c) and (d) illustrate the position and depth errors derived for all inversions of GPO items without distinction of success or failure of the inversion. Targets 13, 17, 22 and 15 stand out with depth errors larger than 25 cm. Their respective depths are 111, 107, 37 and 76 cm. Target 13 also has a large horizontal location error. Inversions for the first three targets were qualified as "failed". Note that EM-61 cart data were in correct GPS coordinates and had not been shifted 11.37 cm to the East and 22.4 cm to the North (the official coordinate system of the study).

Compared to other datasets, inversion of EM-61 cart data generally shows poorer performance. This can be explained by the wide line spacing and lack of attitude information, as opposed to the MTADS data where the coverage was superior and positional errors were limited by the simultaneous recording of several lines at a time. However, all is not lost, as the targets of interest are large and we believe that discrimination of the EM-61 cart data will still result in reduced excavations.

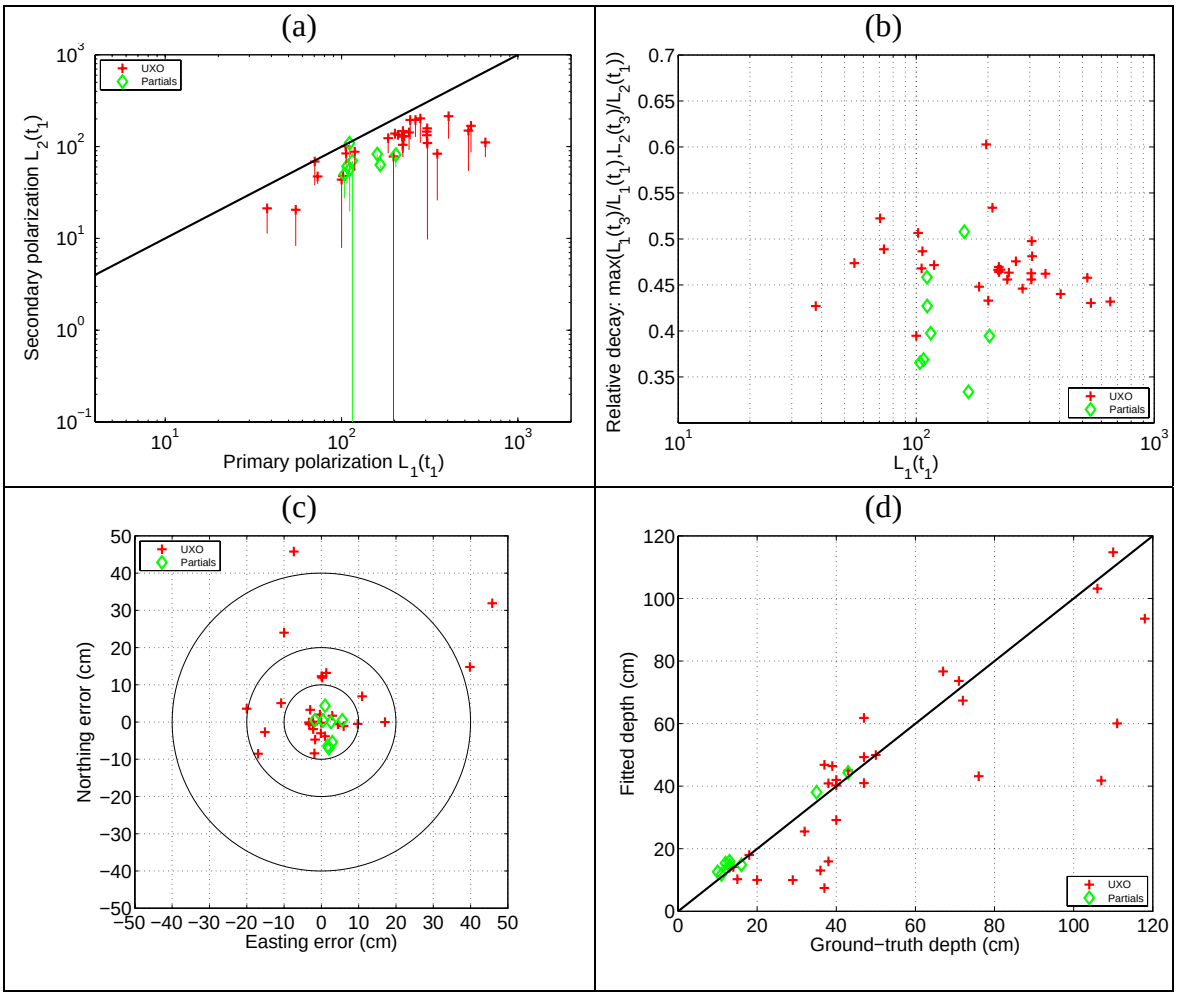



**Figure 5: (a)** Error in estimated dipole locations; **(b)** Predicted versus actual depths.

### b. Analysis of ground-truth data

The rest of the data were inverted using the same procedures as the GPO. There were 633 inverted anomalies, with 413 having acceptable 3 dipole fit and 220 having an unacceptable fit. The large number of rejected fits is a function of the original picks supplied along with the EM-61 cart data. Many of the picks were geologic in nature (particularly in the SW region) and dipole fits were not attainable. The ground-truth for 87 of the EM-61 cart anomalies were provided by the Program Office in early August, with 55 of the anomalies having valid dipole fits. The dipole parameters and ground-truth information are summarized in Figure 6.

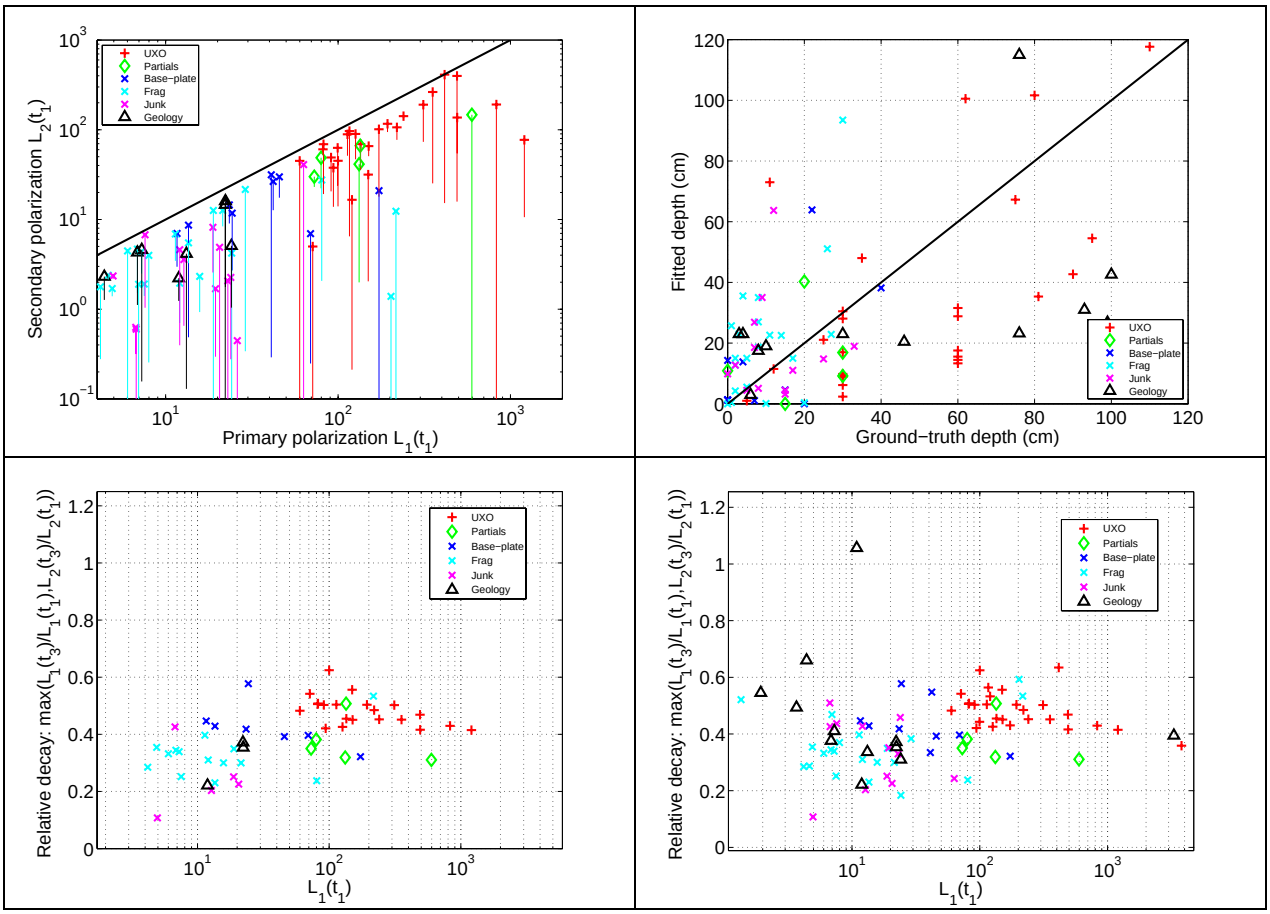



**Figure 6: (a)** Comparison of primary polarization at the first time channel to the secondary polarizations. The extent of the vertical lines are defined by $L_2(t_1)$ and $L_3(t_1)$; **(b)** Comparison of the primary and secondary polarization decay rates (i.e. the ratio between the third and first time channel); **(c)** Comparison of the relative polarization decay with amplitude of the polarizations at the first time channels, **(d)** Repeating the plot in (c) except all fits - including failed fits - are considered.

Figure 6b shows that depths are generally poorly resolved, especially for deeply buried targets. The depth uncertainty directly affects the amplitude of the polarization parameters because there is a trade off between the size and the depth of a buried object. That issue clearly appears in Figure 6a with a wide range of recovered primary polarizations for the 4.2' mortar. Secondary polarizations, which should have similar magnitude for a body of revolution such as a mortar, also show large differences (vertical error bars spanning several orders of magnitude). As with

the GPO, the UXO can be distinguished from a lot of the clutter by virtue of large size and slow decay rate (Figures 6c and 6d).

### c. Creation of a dig-list for the EM-61 cart

For the EM-61 cart data we will produce two dig-sheets: (i) based on a size parameter and (ii) using statistical classification of derived features. Between the GPO and ground-truth data it is apparent that there is significant overlap (in terms of size and decay rate) with the partial rounds. In addition, the large number of "failed" fits would result in a significant number of excavations under the "can't model" category. We therefore investigated methods to reduce the number of items that need to be excavated while hopefully maintaining the same probability of correct classification.

Close scrutiny of failed inversions and inversions that yielded inaccurate depth estimates on the GPO and ground truth targets revealed that poor inversions could be tied to certain features of the data or the inversion. This motivated us to identify and establish rules to define a confidence factor for a given inversion. We consider several criteria that we gather under a so-called Figure of Merit (FOM), which comprises:

- Data features:
    - Signal to noise ratio (SNR): SNR should be above a given threshold for reliable inversion of each time channel, SNR should decay with time if the sensor operates properly and noise estimates are accurate;
    - Data coverage of anomaly: coverage should sample the spatial decay of the EM scattered field to allow recovery of orthogonal polarizations;
- Inversion features:
    - Quality of fit: misfit, correlation coefficient;
    - Variance of estimated depth: there can be several solutions of the inverse problem with similar misfits but distributed over a large range of depth.

The FOM analysis can be performed both on passed and failed inversions. For instance, consider Figure 7. Target 17 placed at 1.07 m depth on the GPO illustrates several of the aspects of a low FOM: the data coverage is poor with wide line spacing over the peak amplitude region; the depth-misfit curve in the upper right corner shows a wide range of depths for models with similar misfits. Therefore it is not surprising to find that the recovered polarizations are outliers in model space; this target has the lowest primary polarization of all UXO because its depth is off by 0.65 m.

Thresholds and relationships can be defined for the parameters intervening in the FOM (systematic tests are currently under way to establish a robust framework) so that a single FOM figure can be assigned to each inversion to reflect the level of confidence in the recovered model. For this demonstration, we elected to use a binary FOM with FOM = 1 indicating a good fit and FOM = 0 a bad-fit. The FOM=0 value was assigned whenever any of the four metrics above exceeded a specified threshold value. These threshold values were assigned by investigating inversions that had poor agreement with the ground-truth position or depth.

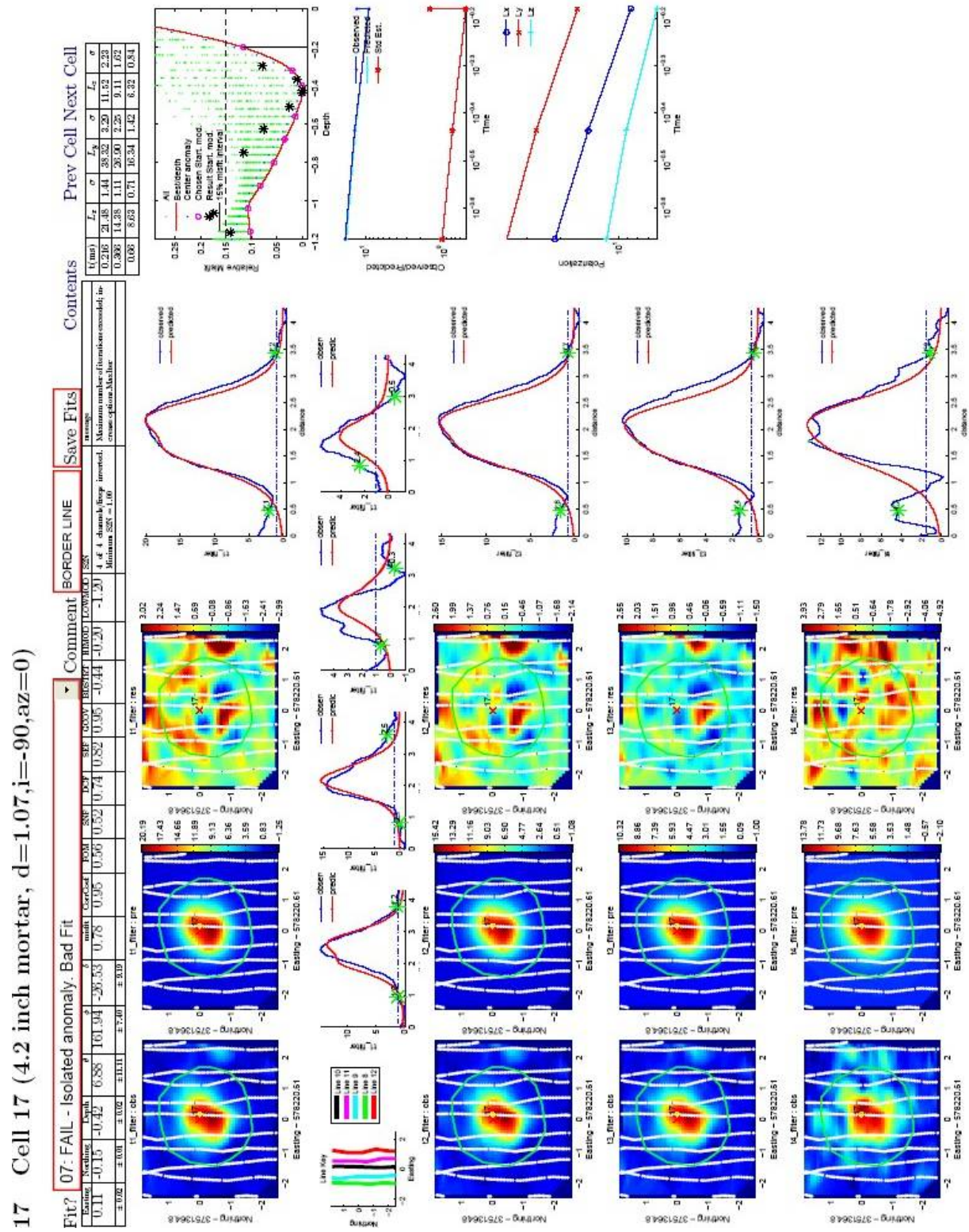



**Figure 7:** Illustration of how poor data coverage and a poor fit contribute to a low FOM value.

Figure 8 plots the feature vectors we will use for classification along with a dot for items with FOM = 0 (unreliable fits) and a circle for items whose fit's were failed by the analyst. Close inspection of the UXO class reveals that the high FOM anomalies are much more tightly clustered than the low FOM anomalies. This suggests that we could set a more aggressive classification threshold for high FOM anomalies compared to low FOM anomalies. We therefore elected to set different thresholds (for both the size and classification based methods) for the low and high FOM anomalies. We also use the feature vectors of failed fits to establish this ranking (as long as the item was failed due to an imperfect fit as opposed to a data-glitch or due to insufficient data: e.g. as may happen at the edge of a grid).

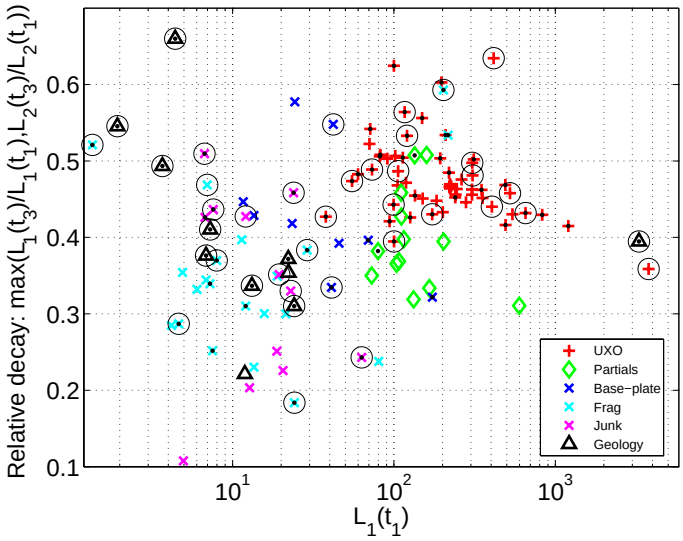



**Figure 8:** Feature space used for inversion that includes a black dot to indicate FOM = 0 and a black circle for failed inversions.

For the size-based dig-sheet we will use $L_1(t_1)$ and we find that all UXO with FOM=1 have $L_1(t_1)>70$, and all UXO with FOM=0 have $L_1(t_1) > 35$. To be conservative, we use the lowest observed $L_1(t_1)$ value at FOM=0 to set the threshold for FOM = 1, and half of that same value to set the FOM=0 threshold (see Table 1). We will place the "can't analyze" category in between the "can't make a decision" categories for FOM = 0 and FOM = 1 (with the FOM = 1 anomalies closest to the high-probability non-UXO class).

For the statistical classification we trained a Probabilistic Neural Network classifier on all "passed" feature vectors from the GPO and ground-truth datasets. We used $L_1(t_1)$ and the maximum of $L_1(t_1)/ L_3(t_1)$ and $L_2(t_3)/ L_2(t_1)$ for the classification. The two feature vectors were first standardized so that they had zero mean and unit standardization. All non-UXO items were combined into a single class. The resulting classifier appears to be intuitively reasonable (Figure 9). Small items with fast time-decays are highly unlikely to be UXO and can be safely left in the ground. At a given size-value, the slower the time decay the more likely the item is a UXO.

The dig-sheet will be ordered according to $P_{Non\text{-}UXO}$ (the PNN probability the item belongs to the non-UXO class) with the item *least-likely* to be a UXO appearing first. We then need to select threshold values of $P_{Non\text{-}UXO}$ for both the high and low FOM anomalies. For FOM = 1 we can be more aggressive in our selection of the threshold so that we can stop digging high FOM anomalies sooner than low FOM anomalies. By investigation of the classification contours we

selected the cut-off values shown in Table 2. Note that the absolute values of the PNN probability are dependent on our (heuristic) choice of smoothing kernel width. We therefore recommend that regulators pay more attention to the digging order (which is only weakly dependant on the kernel width) and stop-digging point than to the specific values of the cut-off probabilities.

| **Dig-sheet category** | **$L_1(t_1)$ threshold (FOM = 0)** | **$L_1(t_1)$ threshold (FOM = 1)** |
|---|---|---|
| High probability non-UXO (don't dig) | $L_1(t_1) < 17$ | $L_1(t_1) < 35$ |
| Can't make a decision (dig): | $17 < L_1(t_1) < 70$ | $35 < L_1(t_1) < 70$ |
| High probability UXO (dig) | $L_1(t_1) > 70$ | $L_1(t_1) > 70$ |

**Table 1:** Thresholds for dig-sheet creation based on the size of the polarization recovered from EM-61 cart data.

| **Dig-sheet category** | **$P_{Non\text{-}UXO}$ threshold (FOM = 0)** | **$P_{Non\text{-}UXO}$ threshold (FOM = 1)** |
|---|---|---|
| High probability non-UXO (don't dig) | $P_{Non\text{-}UXO} > 0.9$ | $P_{Non\text{-}UXO} > 0.75$ |
| Can't make a decision (dig): | $0.5 < P_{Non\text{-}UXO} < 0.9$ | $0.45 < P_{Non\text{-}UXO} < 0.75$ |
| High probability UXO (dig) | $P_{Non\text{-}UXO} < 0.5$ | $P_{Non\text{-}UXO} < 0.45$ |

**Table 2:** Thresholds for dig-sheet creation based on the PNN classification of the EM-61 cart-data.

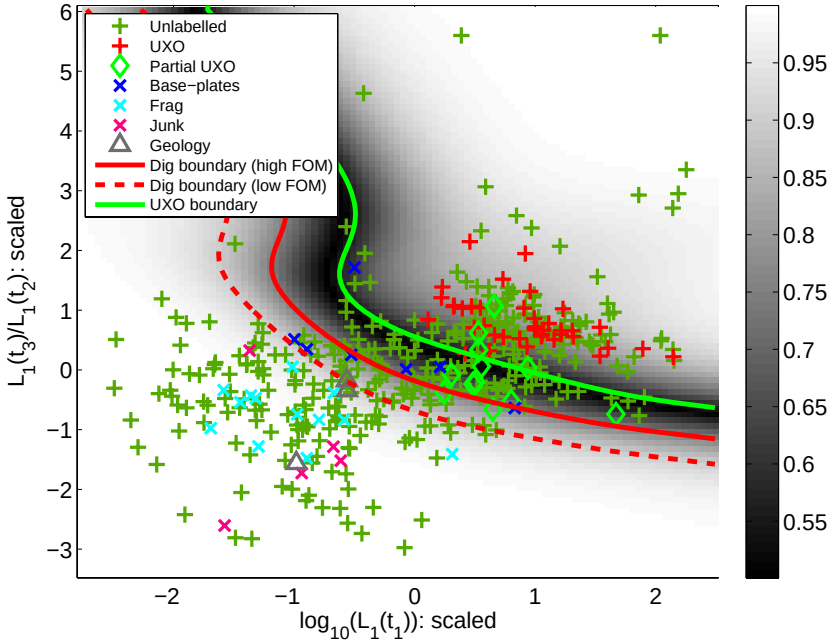



**Figure 9:** PNN classifier trained on $\log_{10}(L_1(t_1))$ and the maximum of $L_1(t_1)/ L_3(t_1)$ and $L_2(t_3)/ L_2(t_1)$. Both labeled and unlabelled feature vectors are shown. The color-scale is constructed such that black corresponds to equal probability of membership to either class. The grayscale inside/outside the UXO/non-UXO region indicates increasing probability of membership of the UXO/non-UXO class. Green boundary delineates the high confidence UXO class, the solid red line is the high FOM non-UXO boundary and the dashed red line is the low FOM boundary.

# 3. MTADS EM-61 ARRAY DATA

## a. Analysis of Geophysical Prove Out results

Processing and inversion parameters for the MTADS EM-61 data from the GPO were the same as those for the EM-61 cart (including the estimated 10% error term).

The instantaneous amplitude 3 dipole model fits were found to be acceptable for all test-plot items with the exception of targets 10 and 14 (1.1 and 1.18 m depths respectively). Figure 10 plots different combinations of the recovered polarization parameters. Figure 10 (a) plots the primary and secondary polarizations in a similar manner to Figure 5a for the EM-61 cart. The data for four of the anomalies were unable to constrain the secondary polarizations of the target, suggesting that size will be a better discriminant than the spread in the secondary polarizations. The limited discrimination of the secondary polarization spread is confirmed in Figure 10 (b).

A comparison of the relative decay rates for the primary and secondary polarizations are plotted in 10 (c). This combination of parameters also produces good separation between UXO and non-UXO targets, with UXO targets having, in general, a slower decay rate.

The results in Figure 10(a)-(c) suggest that good separation between UXO and non-UXO targets can be achieved through a combination of primary polarization decay and polarization size. Figure 10 (d) plots a measure of size (i.e., the product of the primary and secondary polarizations at the first time channel) and a measure of the decay (the larger of the secondary and primary polarization decay rates). This combination of parameters produces the best separation between UXO and non-UXO targets.

Figure 11(a) illustrates the errors in northing and easting positions, calculated as the difference between the positions reported by the fitted depth and the supplied ground truth positions. The fitted depths obtained via inversions are plotted versus the ground truth depths in Figure 11(b). The 1:1 line indicates where the partial and intact 4.2” inch mortars would plot were the fitted and ground truth values in perfect agreement. All but four of the items contain positioning errors of less than 10 cm, with only two items exhibiting positioning errors greater than 20 cm. All predicted depths are within 25 cm of the ground truth values, and thirty of the thirty-six items with valid fits are within 10 cm.

Contrast these results for the MTADS EM61 data with those presented in Figure 5 for the EM-61 cart. The improved data quality, positional and orientation integrity of the MTADS EM61 certainly pays off in terms of better separation between intact and partial 4.2” mortars.

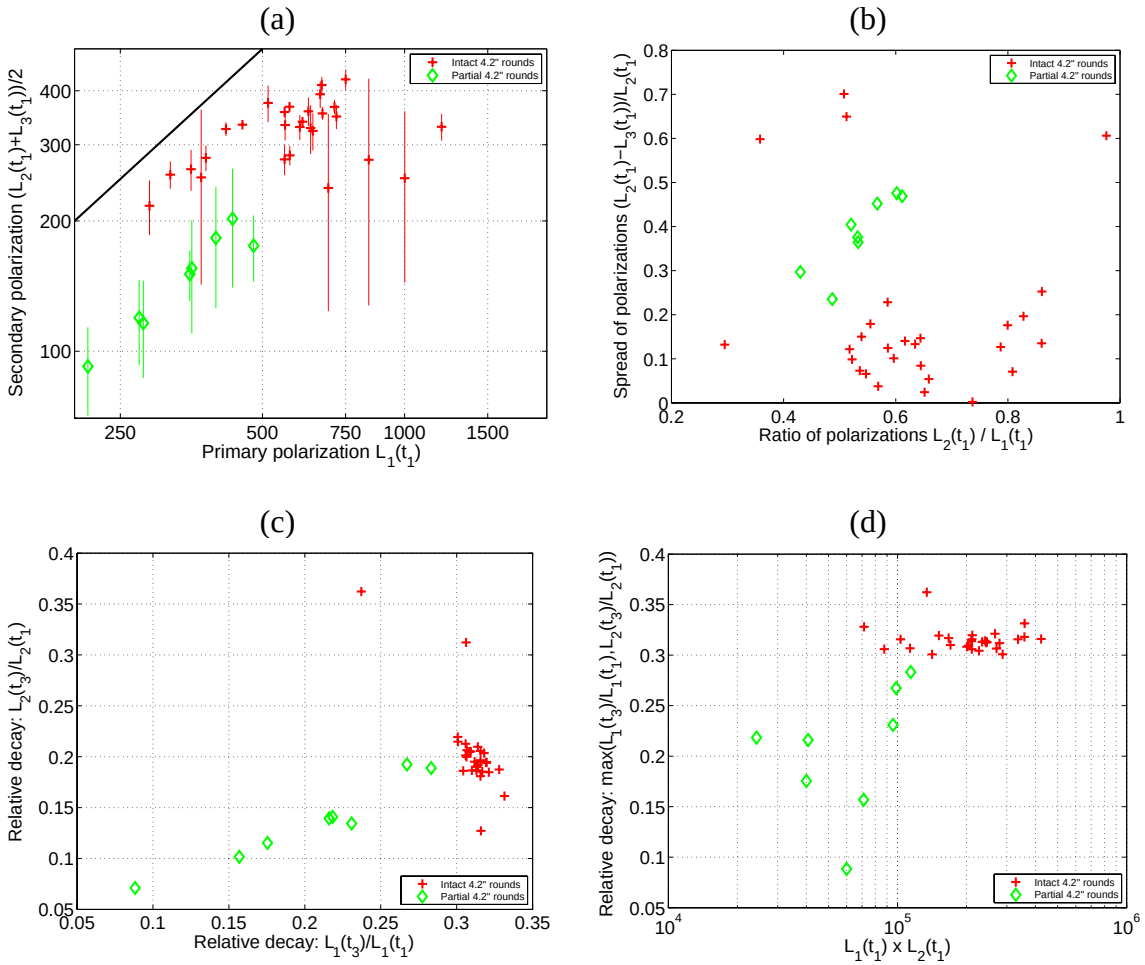



**Figure 10: (a)** Comparison of primary polarization at the first time channel to the secondary polarizations. The extent of the vertical lines are defined by $L_2(t_1)$ and $L_3(t_1)$; **(b)** Comparison of the primary and secondary polarization decay rates (i.e. the ratio between the third and first time channel); **(c)** Comparison of the relative polarization decay with amplitude of the polarizations at the first time channels, **(d)** Repeating the plot in (c) except all fits – including failed ones - are considered.

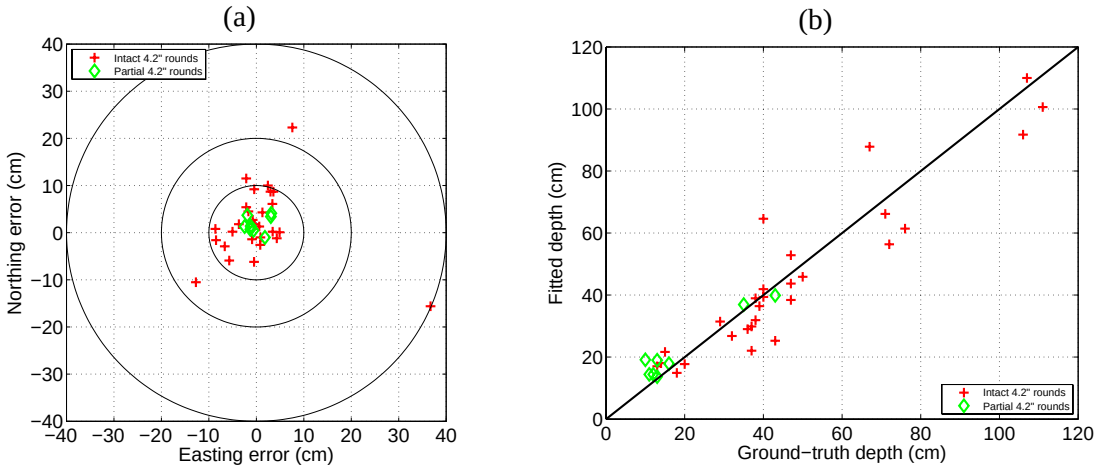



**Figure 11: (a)** Error in estimated dipole locations; **(b)** Predicted versus actual depths.

### b. Analysis of ground-truth data

The rest of the data were inverted using the same procedures as the GPO. There were 870 inverted anomalies, with 414 having acceptable 3 dipole fit and 456 having an unacceptable fit. The large number of rejected fits is a function of the original picks supplied along with the MTADS EM-61 array data. Many of the picks were geologic in nature (particularly in the SW region) and good dipole fits were not attainable. The ground-truth for 136 of the MTADS EM-61 array anomalies was provided by the Program Office in early August, with 59 of the anomalies having valid dipole fits. Twenty-seven of the ground-truth items were from seeded 4.2" mortars (there were two additional deep rounds in the ground-truth that were not detected, 804 @ 75cm, 1058 @ 120cm). The dipole parameters and ground-truth information are summarized in Figure 12.

Figure 12 summarizes the recovered dipole polarizations for the ground truth data. Figure 12(a) plots the primary polarization at the first time channel (i.e. $L_1(t_1)$) and the secondary polarizations at the first time channel (i.e. $L_2(t_1)$ and $L_3(t_1)$). If we focus on the red crosses in Figure 11(a), it is clear that (a) there is good separation between the 4.2 inch mortars and the non-UXO targets (with the exception of two of the partial mortars) using this size measure, and (b) the data for three of the 4.2 inch mortar anomalies were unable to constrain the secondary polarizations. A comparison of the relative decay rates for the primary and secondary polarizations are plotted in Figure 12(b). The relative decay is defined as the ratio of the polarizations at the 3$^{rd}$ and 1$^{st}$ time channels. This combination of parameters also produces good separation between UXO and non-UXO targets. Figure 12(c) plots a measure of size (i.e., the product of the primary and secondary polarizations at the first time channel) and a measure of the decay (the larger of the secondary and primary polarization decay rates). As was shown in Figure 10, this combination of parameters produces a clear separation between UXO and non-UXO targets.

For each of the parameter plots with passed inversions (i.e., panels (a), (b) and (c)) three of the 4.2 inch mortar anomalies appear to be outliers from the cluster in model space. Figure 12(d) repeats the parameters plotted in Figure 12(c) except all inversions, including those classified as failed fits, are included. Surprisingly, a significant number of picks were due to geology.

Figures 13(a) and (b) show that the recovered locations and depths agree very well with the reported ground-truth locations. Note that there is a tendency to predict shallower depths for the cluster of seed items at 60 cm. We suspect the items were actually buried a little shallower as the magnetics and other EM methods show a similar bias.

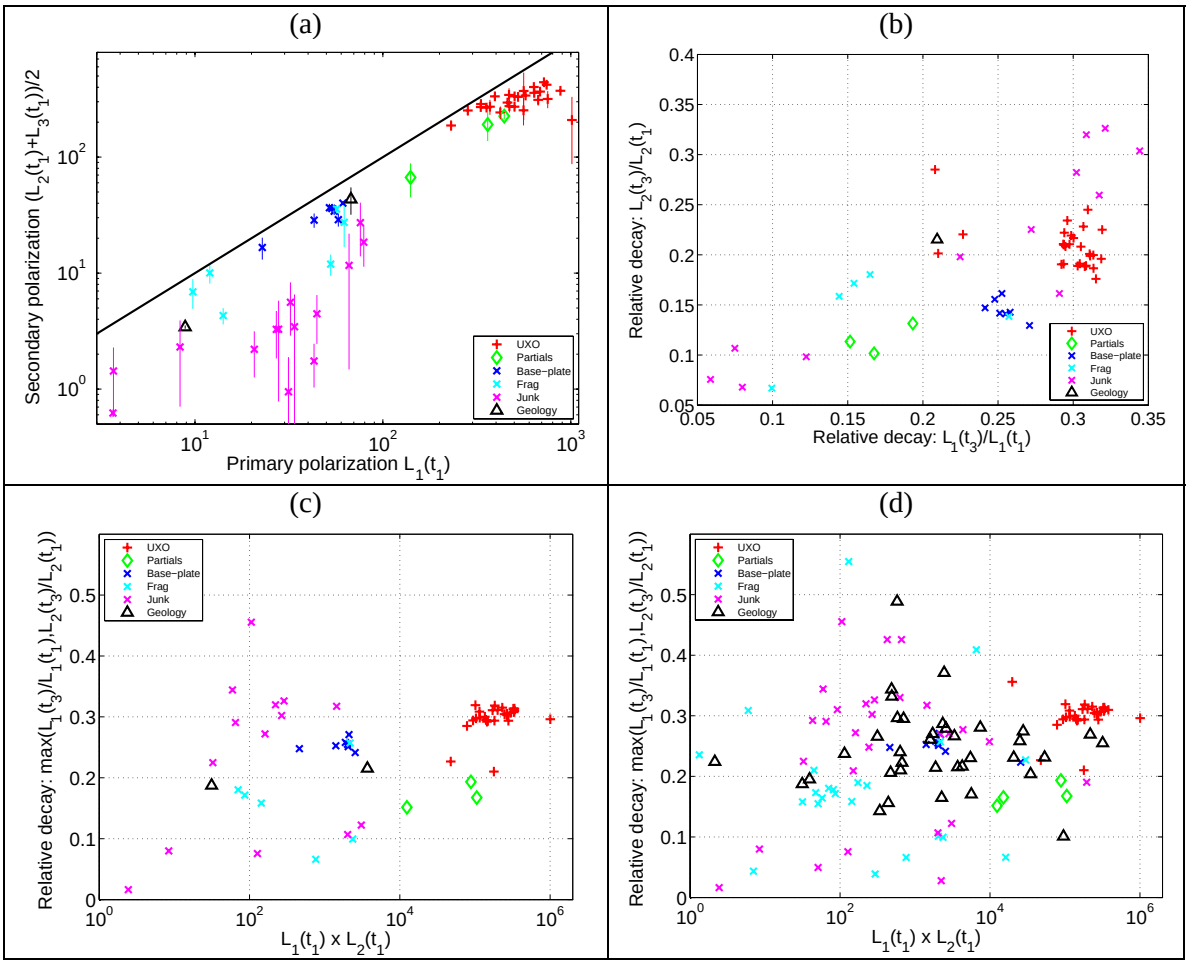



**Figure 12: (a)** Comparison of primary polarization at the first time channel to the secondary polarizations. The extent of the vertical lines are defined by $L_2(t_1)$ and $L_3(t_1)$; **(b)** Comparison of the primary and secondary polarization decay rates (i.e. the ratio between the third and first time channel); **(c)** Comparison of the relative polarization decay with amplitude of the polarizations at the first time channels, **(d)** Repeating the plot in (c) except all fits are considered.

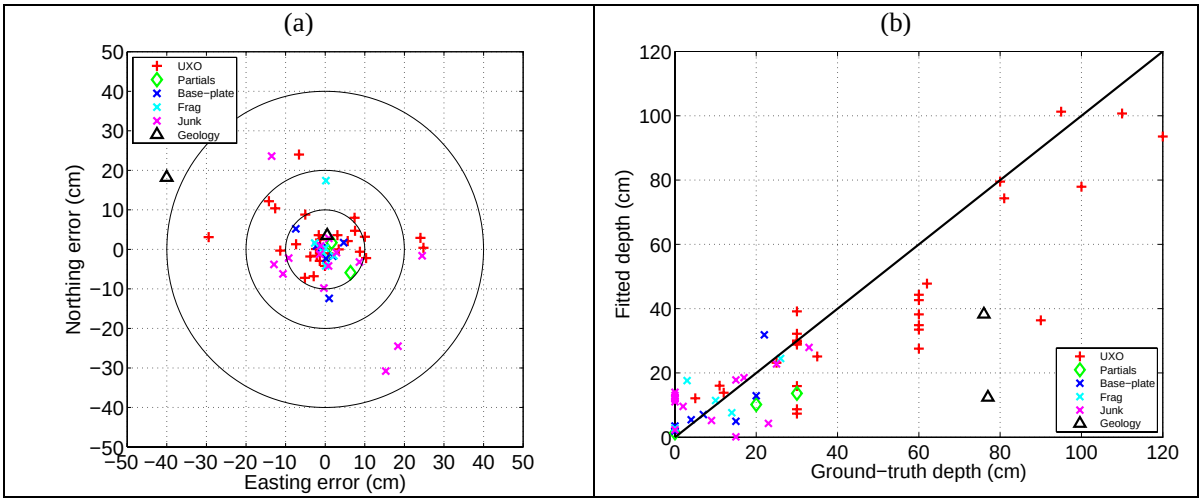



**Figure 13: (a)** Error in estimated dipole locations; **(b)** Predicted versus actual depths

### c. Discrimination strategy for MTADS EM61 data

We will submit two dig-sheets for the MTADS EM61 data (one based on size alone, the other on statistical classification) as well as a third dig-sheet for the MTADS EM61 data interpreted cooperatively (see section 5). We will use the same methodology as for the EM-61 cart data. That is, we will include failed fits and use different thresholds for low and high FOM anomalies. This will be particularly important for the MTADS EM61 data: otherwise the large number of geologic anomalies failed by the analysts will result in a high FAR.For the size-based dig-sheet we will use $L_1(t_1)$ and we find that all UXO with FOM=1 have $L_1(t_1)>680$, and all UXO with FOM=0 have $L_1(t_1) > 220$ (Table 3).

| **Dig-sheet category** | **$L_1(t_1)$ threshold (FOM = 0)** | **$L_1(t_1)$ threshold (FOM = 1)** |
|---|---|---|
| High probability non-UXO (don't dig) | $L_1(t_1) < 110$ | $L_1(t_1) < 220$ |
| Can't make a decision (dig): | $110 < L_1(t_1) < 680$ | $220 < L_1(t_1) < 680$ |
| High probability UXO (dig) | $L_1(t_1) > 680$ | $L_1(t_1) > 680$ |

**Table 3:** Thresholds for dig-sheet creation based on the size of the polarization recovered from the MTADS EM61 data.

For the statistical classification we trained a PNN classifier on $L_1(t_1)$ and the maximum of $L_1(t_1)/ L_3(t_1)$ and $L_2(t_3)/ L_2(t_1)$. Prior to training the feature vectors were standardized to have 0 mean and unit standard deviation (Figure 14a). UXO anomalies with high FOM are well separated from the rest of the items in the ground-truth and an aggressive cut-off could be selected. To prevent false-negatives due to an aggressive strategy we opt to set a conservation cut-off value of $P_{Non\text{-}UXO} = 0.7$ for the high FOM anomalies. Two UXO's with low FOM (Figure 14b) lie close to this boundary we therefore select an even less aggressive cut-off value of 0.9 for the low FOM anomalies. Thresholds for construction of the dig-sheet are listed in Table 4.

| **Dig-sheet category** | **$P_{Non\text{-}UXO}$ threshold (FOM = 0)** | **$P_{Non\text{-}UXO}$ threshold (FOM = 1)** |
|---|---|---|
| High probability non-UXO (don't dig) | $P_{Non\text{-}UXO} > 0.9$ | $P_{Non\text{-}UXO} > 0.7$ |
| Can't make a decision (dig): | $0.5 < P_{Non\text{-}UXO} < 0.9$ | $0.45 < P_{Non\text{-}UXO} < 0.7$ |
| High probability UXO (dig) | $P_{Non\text{-}UXO} < 0.5$ | $P_{Non\text{-}UXO} < 0.45$ |

**Table 4:** Thresholds for dig-sheet creation based on the size of the PNN classifier applied to feature vectors extracted from the MTADS EM61 data.

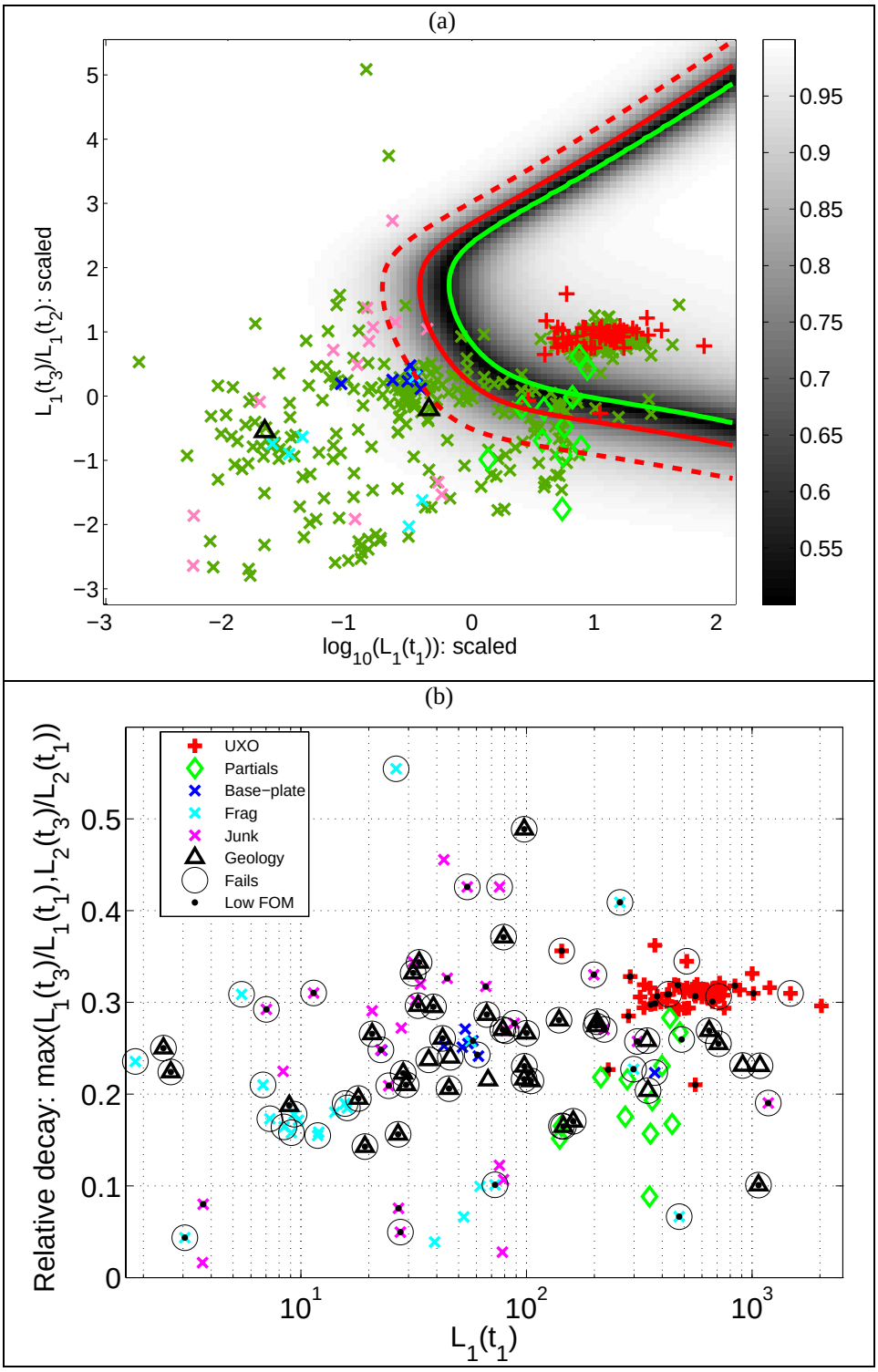



**Figure 14: (a)** PNN classifier trained on $\log_{10}(L_1(t_1))$ and the maximum of $L_1(t_1)/ L_3(t_1)$ and $L_2(t_3)/ L_2(t_1)$ after standardizing the feature vectors. Green boundary delineates the high confidence UXO class, the solid red line is the high FOM non-UXO boundary and the dashed red line is the low FOM boundary. (b) Feature space including failed fits and with low FOM anomalies identified.

# 4. EM-63 DATA

## a. Analysis of Geophysical Prove Out results

The following parameters were used to invert the EM-63 data on the GPO:

- Estimates of data errors:
  - Background noise was estimated by choosing data surrounding the anomaly, and calculating statistics for those data.
  - We assumed 10 percent noise on each data point.
- *Masking*: The spatial extent of data used in the inversions was determined using the elliptical masking technique described in the Demonstration Plan.
- *Model*: The data were fit using 3 unique polarizations for the dipole tensor with

$$L_i(t) = k_i t^{-\beta_i} \exp(-t/\gamma_i)$$

The resulting model vector is:

$$m = \left[X, Y, Z, \phi, \theta, \psi, k_1, \beta_1, \gamma_1, k_2, \beta_2, \gamma_2, k_3, \beta_3, \gamma_3\right]$$

where $(X,Y,Z)$ is the location, $(\phi,\theta,\psi)$ are the orientation angles

The 3 dipole model fits were found to be acceptable for all test-plot items with the exception of five targets (3, 10, 13, 14, 17, all of which were buried at depths greater than 1m). Note that we were not able to use the "pitch lines" due to the significant geological response in the SW area that prevented accurate recovery of dipole parameters (the inversion would require the estimation of a background model simultaneously with the dipole model to use the pitch lines). In addition, even after verifying gains and time-gate information on the top-coil data with Geonics, we were unable to obtain good model fits to the top-coil data. Therefore, we only used the 26 bottom-coil time-channels.

Each panel in Figure 15 plots a combination of the decay parameters for the primary polarization (i.e. $k_1$, $\beta_1$, and $\gamma_1$) and secondary polarizations (i.e. $k_2$ and $k_3$). Figures 15(a) and (b) plot two different measures of the primary and secondary polarization. The polarization amplitude is proportional to the size of the target. Figure 15(a) compares the primary polarization $k_1$ to an average of the two secondary polarizations. The magnitudes of $k_2$ and $k_3$ determine the extent of the lines in the y-direction. Figure 15(b) plots the integral of the polarizations (the area under the Pasion-Oldenburg decay curve). Although the separation between the 4.2 inch mortars and non-UXO targets are similar in Figures 15(a) and (b), using the integral of the polarizations produces somewhat clearer separation. In Figure 15(b), there is one partial UXO target (number 36 in the GPO) that prevents this combination of parameters from producing a distinct separation between the intact and partial UXO. The data for this target was unable to accurately constrain the depth of the target. The estimated depth for this target was too deep, and thus the estimated polarization amplitude is quite large. Figure 15(c) plots the decay parameters β and γ for the primary polarization. The time constant $\gamma$ for non-UXO items is much less than for the 4.2 inch mortar. The parameters β and γ can effectively be combined by using a relative decay rate (polarization at channel $n$ divided by polarization at channel 1). Figure 15 (d) plots $L_1(t_{15})/L_1(t_1)$ versus $L_1(t_{10})/L_1(t_1)$ and demonstrates that combining β and γ also provides good class separation.

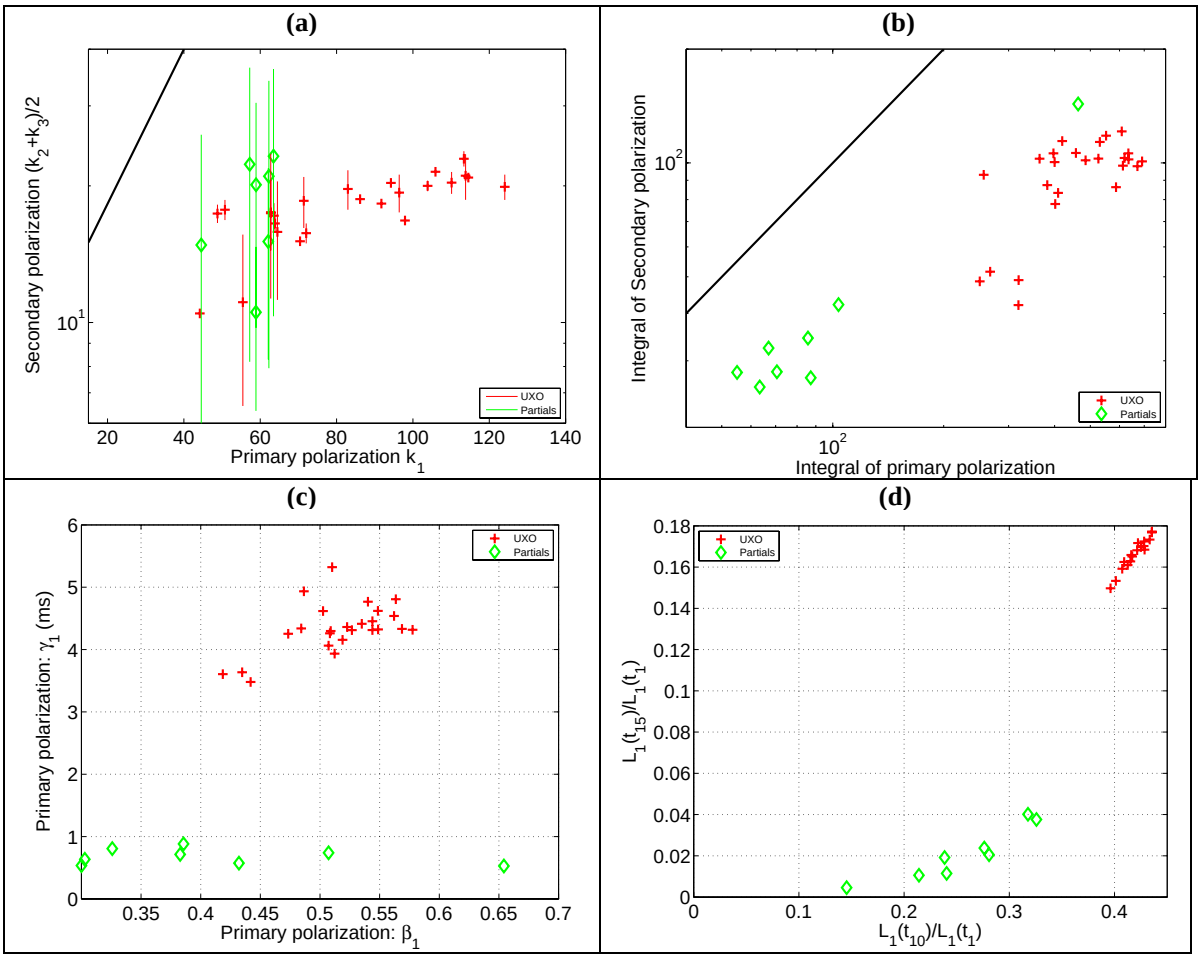



**Figure 15:** Geonics EM-63 inversion parameter plots for the GPO data. Panels **(a)** and **(b)** compare size-based parameters while panels **(c)** and **(d)** plot decay characteristics.

Figure 16 contain plots of the recovered polarizations from the GPO and groundtruth data inversions. In Figure 16(a), we clearly see that one of the partial mortars from the GPO is fit with primary polarization that is quite large. This was due to a poor estimate of the target depth (recovered depth = 0.6 m, true depth = 0.2 m). However, we see that the time constant for the polarization is similar to the other non-UXO items in the GPO. Therefore it should still be possible to discriminate this item from the intact mortars, provided that we use a discriminant that includes the estimated time constant information. This observation is consistent with time constant plot of Figure 15(c). Figures 15 and 16 suggests that a combination of the polarization size and the time constant parameters for the primary polarization has discrimination potential.

Figure 17(a) illustrates the errors in positions, calculated as the difference between the positions reported by the fitted depth and the supplied ground truth positions. The fitted depths obtained via inversions are plotted versus the ground truth depths in Figure 17(b). All but five of the items contain positioning errors of less than 10 cm with only a single item exhibiting errors greater than 20 cm in the positions. All but one of the predicted depths are within 20 cm of the ground truth values and twenty-nine of the thirty-three items with valid fits are within 10 cm. Note that there was insufficient SNR to obtain good fits to any of the deep UXO, exposing a potential weakness of the EM-63.

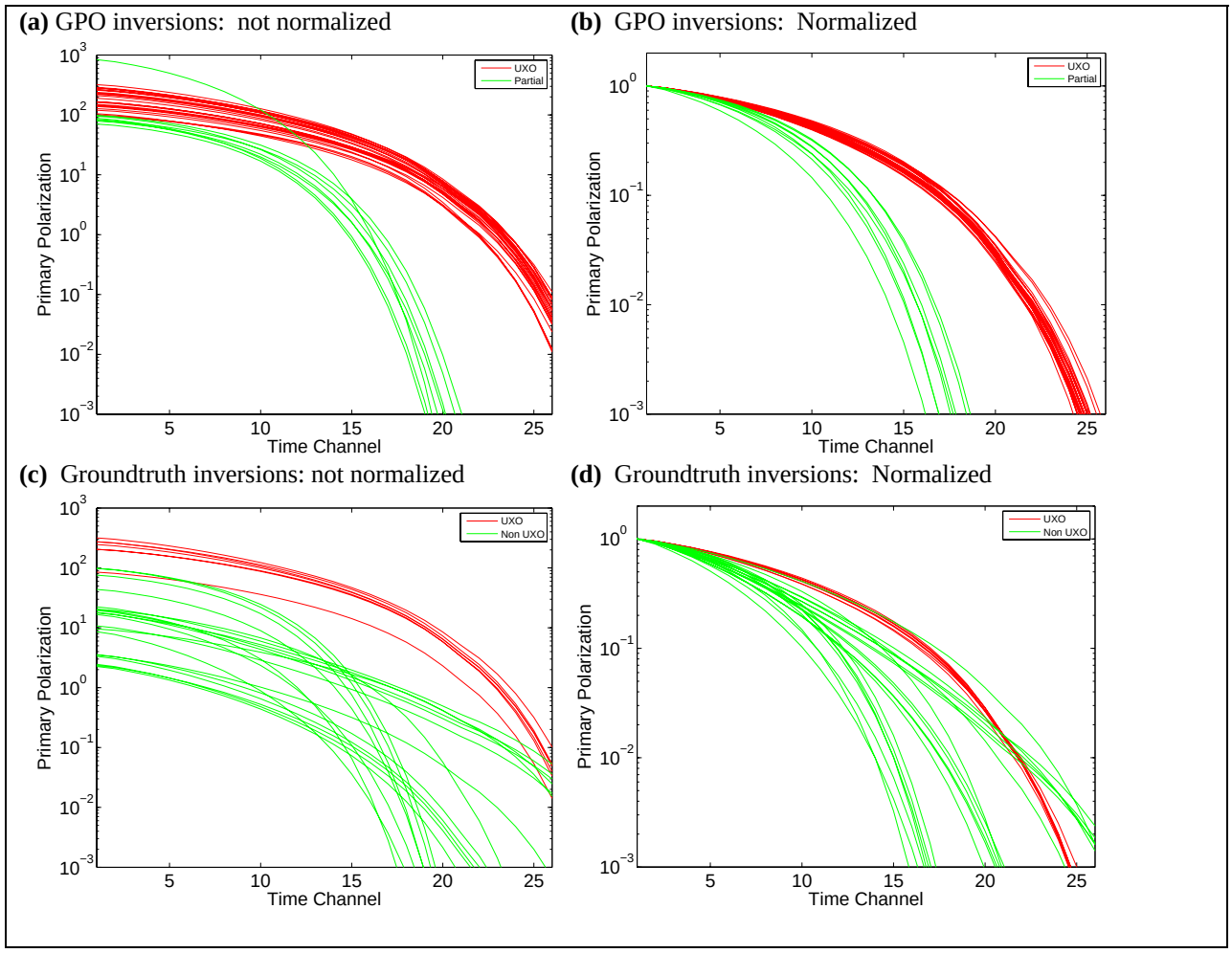



**Figure 16**. Recovered primary polarizations from EM-63 inversions. Panels **(a)** and **(b)** contain raw and normalized recovered primary polarizations from the GPO data, and **(c)** and **(d)** contain raw and normalized recovered primary polarizations from the groundtruth data.

### b. Analysis of ground-truth data

The rest of the data were inverted using the same inversion parameters (i.e. noise and masking parameters) as the GPO. There were 179 inverted anomalies, with 146 having acceptable dipole fit and 33 having an unacceptable fit. The ground-truth for 28 of the EM-63 anomalies was provided by the Program Office in early August, with 27 of the anomalies having valid dipole fits. Eight of the ground-truth items were 4.2" mortars.

The dipole parameters and ground-truth information are summarized in Figure 18. Recovered parameters from three of the UXO appear to fall outside the cluster. Of those three UXO targets, the data are unable to constrain the secondary polarizations to be equal. There is, however, clear separation between UXO and non-UXO when using the polarization magnitudes. Integrals of the polarizations produce an even better separation, with 3 of the targets again occurring outside the cluster. A plot of $\beta_1$ versus $\gamma_1$ (Figure 18c) and parameters derived from $\beta_1$ and $\gamma_1$ (Figure 18d) shows that there is good class separation when using the decay characteristics of the primary polarization. There separation is better with the latter plot and hence we use $L_1(t_{15})/L_1(t_1)$ for classification.

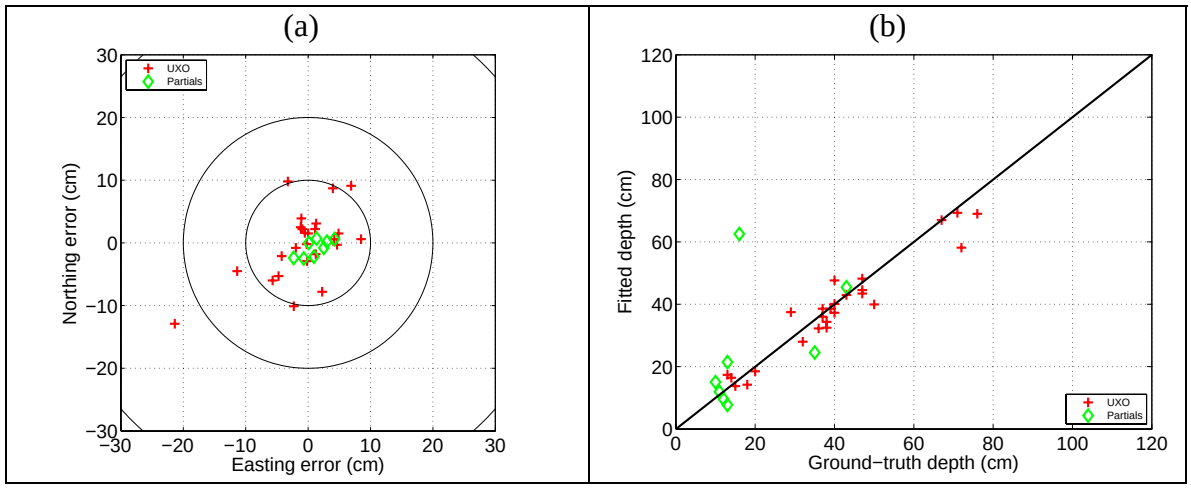



**Figure 17: (a)** Error in estimated dipole locations; **(b)** Predicted versus actual depths.

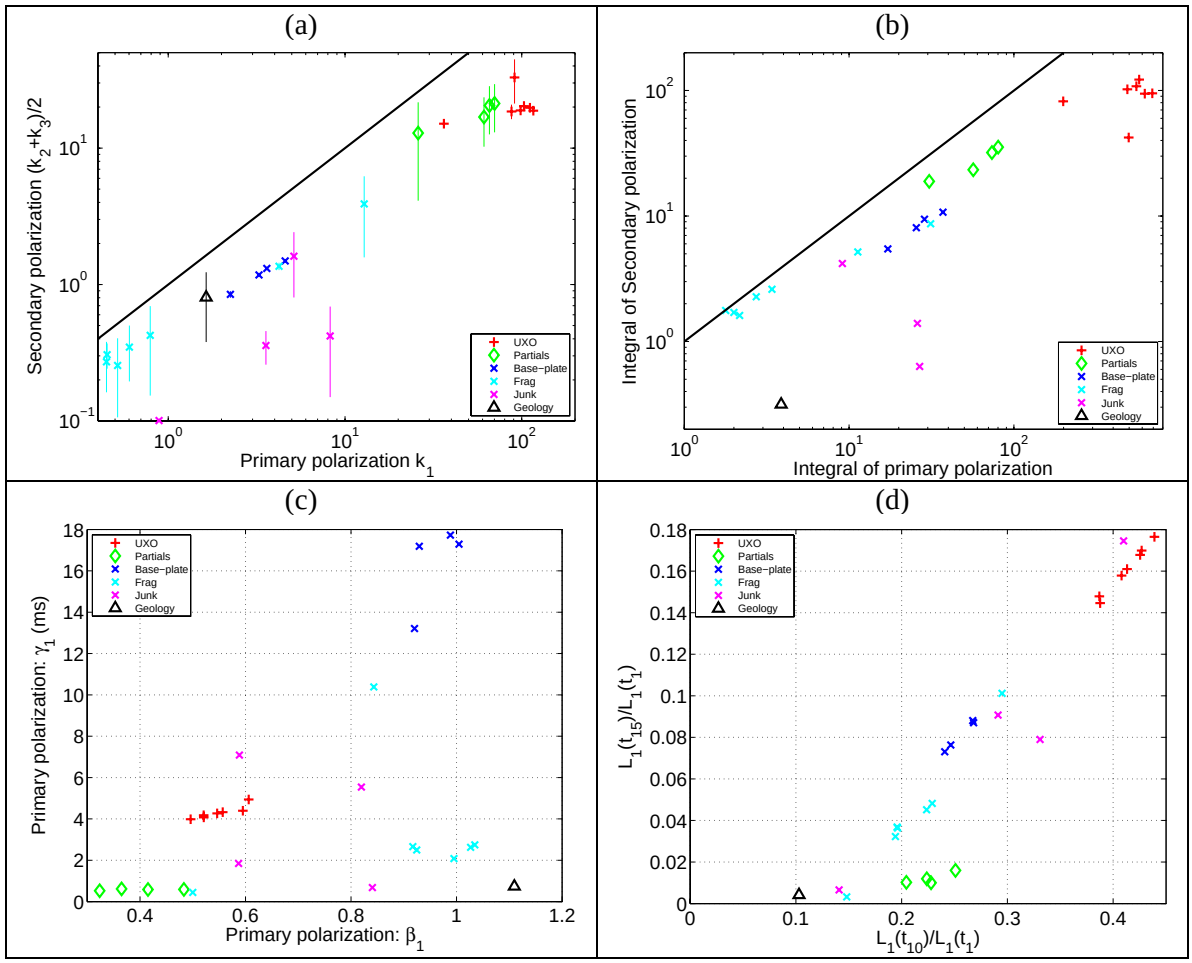



**Figure 18:** Geonics EM63 inversion parameter plots for the non-GPO data with ground truth. Panels **(a)** and **(b)** compare size-type parameters, and panels **(c)** and **(d)** plot the decay characteristics of the primary and polarization.

Figures 16(c) and (d) plot the normalized primary polarizations for the ground truth data. Panel (d) shows that, in contrast to the items in the GPO, the groundtruth plots contain non-UXO items that have similar time constants. The size of the polarization should be used to help discriminate these items.

### c. Discrimination

Analysis of the GPO and the ground-truth data indicates that a combination of a size-based and a decay-based parameter is sufficient to discriminate the 4.2" mortars from the non-hazardous items. The viewpoint we take is that if an item is too small, or has a time-decay parameter that deviates from that of a 4.2" mortar it is extremely unlikely to be an item of concern. Figure 19 shows a PNN classifier with a Gaussian kernel function that was trained on the $\log_{10}(k_1)$ and $L_1(t_{15})/L_1(t_1)$ features from the GPO and ground-truth data. All non-UXO items were combined into a single class prior to training. There is very clear separation between the intact UXO items and the other items with ground-truth. There is a cluster of unlabelled feature vectors inside the support plane for the UXO class that have a high likelihood of being UXO. There are three unlabelled features vectors outside of the main cluster but still within the area where the UXO probability is greater than the non-UXO probability. .

The dig sheet will be ranked by the PNN classifer and we will recommend that all items with $P_{non\text{-}UXO} > 0.9$ be left in the ground. The region between $0.1 > P_{non\text{-}UXO} > 0.9$ (between the red and green lines in Figure 19) will then follow and will go under the designation, "can't make a decision (dig)" . The 33 unlabelled feature vectors with failed fits will then follow with the category "can't analyze (dig)". The items with $P_{non\text{-}UXO} < 0.1$ will then follow and will be classified as high confidence UXO. We will also not use a FOM metric. Note that we believe all the UXO will be in the high confidence UXO or "can't analyze (dig) categories. The reason we suggest digging further down the priority order is to have a safety net to catch items with poor fits.

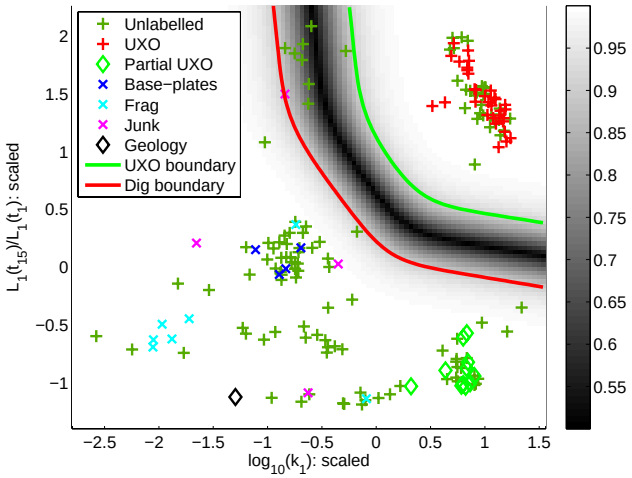



**Figure 19**: PNN classifier applied to $\log_{10}(k_1)$ and $L_1(t_{15})/L_1(t_1)$. The color-scale is constructed such that black corresponds to equal probability of membership to either class. The grayscale inside/outside the UXO/non-UXO region indicates increasing probability of membership of the UXO/non-UXO class. Green boundary delineates the region that is high confidence UXO. Red boundary indicates region of non-UXO.

# 5. MTADS EM-61 ARRAY DATA: COOPERATIVE INVERSION

## a. Analysis of Geophysical Prove Out results

The MTADS EM-61 data set were cooperatively inverted by using dipole location estimates from MTADS magnetics data as *a priori* information. Upper and lower constraints on the location are defined to be twice the estimated variances of the estimated location parameters, i.e.:

$$X^{mag} - 2\sigma_X^{mag} < X < X^{mag} + 2\sigma_X^{mag}$$
$$Y^{mag} - 2\sigma_Y^{mag} < Y < Y^{mag} + 2\sigma_Y^{mag}$$
$$Z^{mag} - 2\sigma_Z^{mag} < Z < Z^{mag} + 2\sigma_Z^{mag}$$

where the estimated location from the inversion of magnetics data is ($X^{mag}$, $Y^{mag}$, $Z^{mag}$) and their estimated standard deviations are $\left(\sigma_X^{mag}, \sigma_Y^{mag}, \sigma_Z^{mag}\right)$. The noise and mask definitions are the same as for the non-cooperatively inverted data.

The cooperative instantaneous amplitude, three-polarization model fits were found to be acceptable for all cooperatively inverted test-plot items with the exception of six targets (3, 10, 13, 14, 23, 32, with four of these six failed fits occurring at depths greater than 1 m). Figure 20 plots combinations of recovered parameters for the cooperatively inverted MTADS EM-61 data collected over the GPO. Figure 20(a) indicates that the data for a number of the anomalies were unable to constrain the secondary polarizations of the target, suggesting that size might be a better discriminant than the spread in the secondary polarizations. The limited discrimination of the secondary polarization spread is confirmed in Figure 20(b). A comparison of the relative decay rates for the primary and secondary polarizations are plotted in Figure 20(c).

Both the amplitude and time-decay based parameters of the 4.2" mortars are more tightly clustered for the cooperatively inverted data (Figure 20) compared to the non-cooperatively inverted data (Figure 10). This is particularly apparent in Figures 10(d) and 20(d). The improved estimates of parameters are due to the more accurate location and depth estimates returned by the cooperative inversion process (compare Figures 11(a) and (b) to Figures 20(e) and (f)). Only 2 items have depth errors of greater than 10 cm for the cooperatively inverted data, compared to 4 items for the non-cooperatively inverted data. The largest error in depth is 18 cm for cooperative compared to just under 25 cm for the non-cooperatively inverted data.

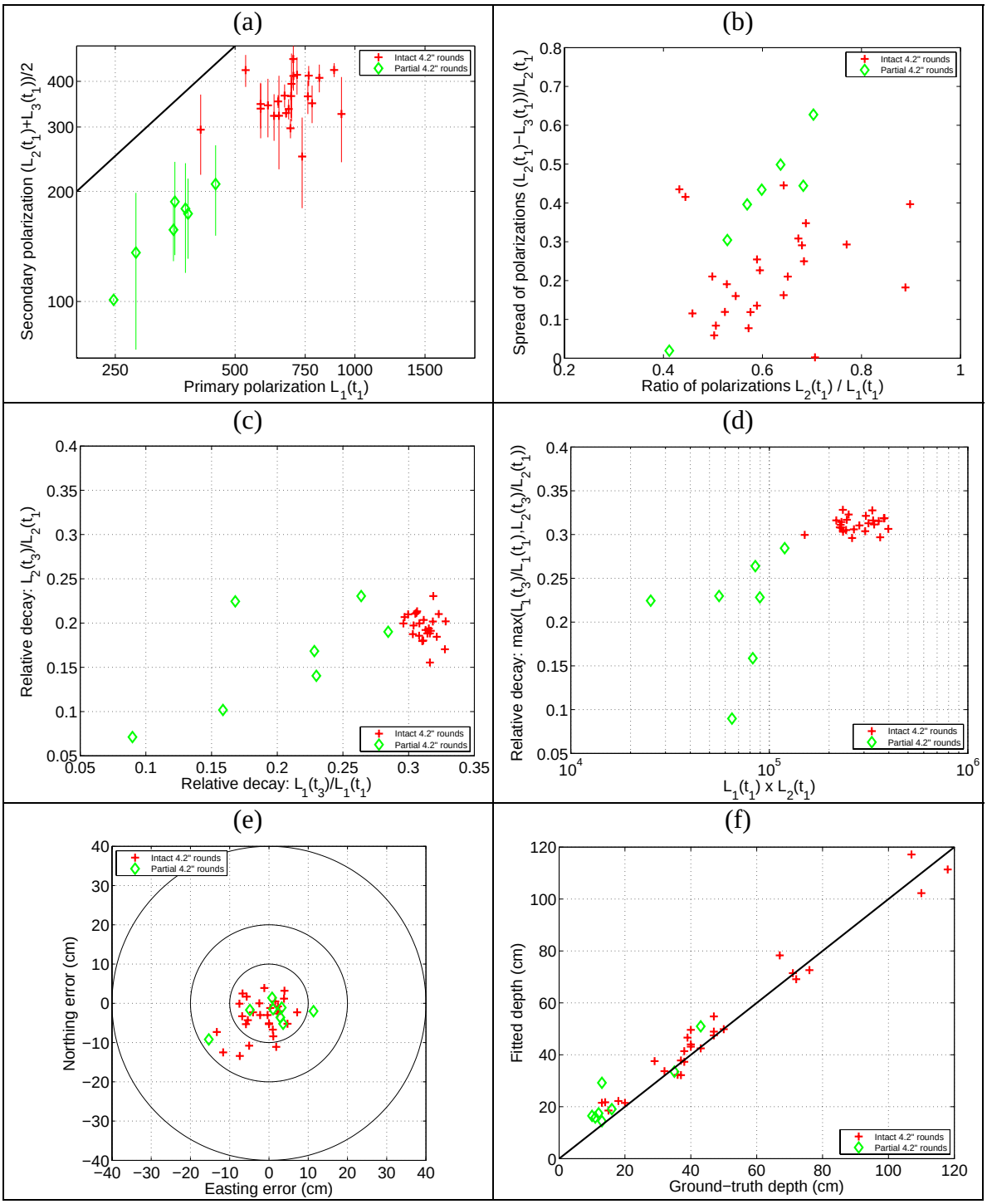



**Figure 20: (a)** Comparison of primary polarization at the first time channel to the secondary polarizations. The extent of the vertical lines are defined by $L_2(t_1)$ and $L_3(t_1)$; **(b)** Comparison of the primary and secondary polarization decay rates (i.e. the ratio between the third and first time channel); **(c)** Comparison of the relative polarization decay with amplitude of the polarizations at the first time channels, **(d)** Repeating the plot in (c) except all fits - including failed fits - are considered. **(e)** Error in estimated dipole locations; **(f)** Predicted versus actual depths.

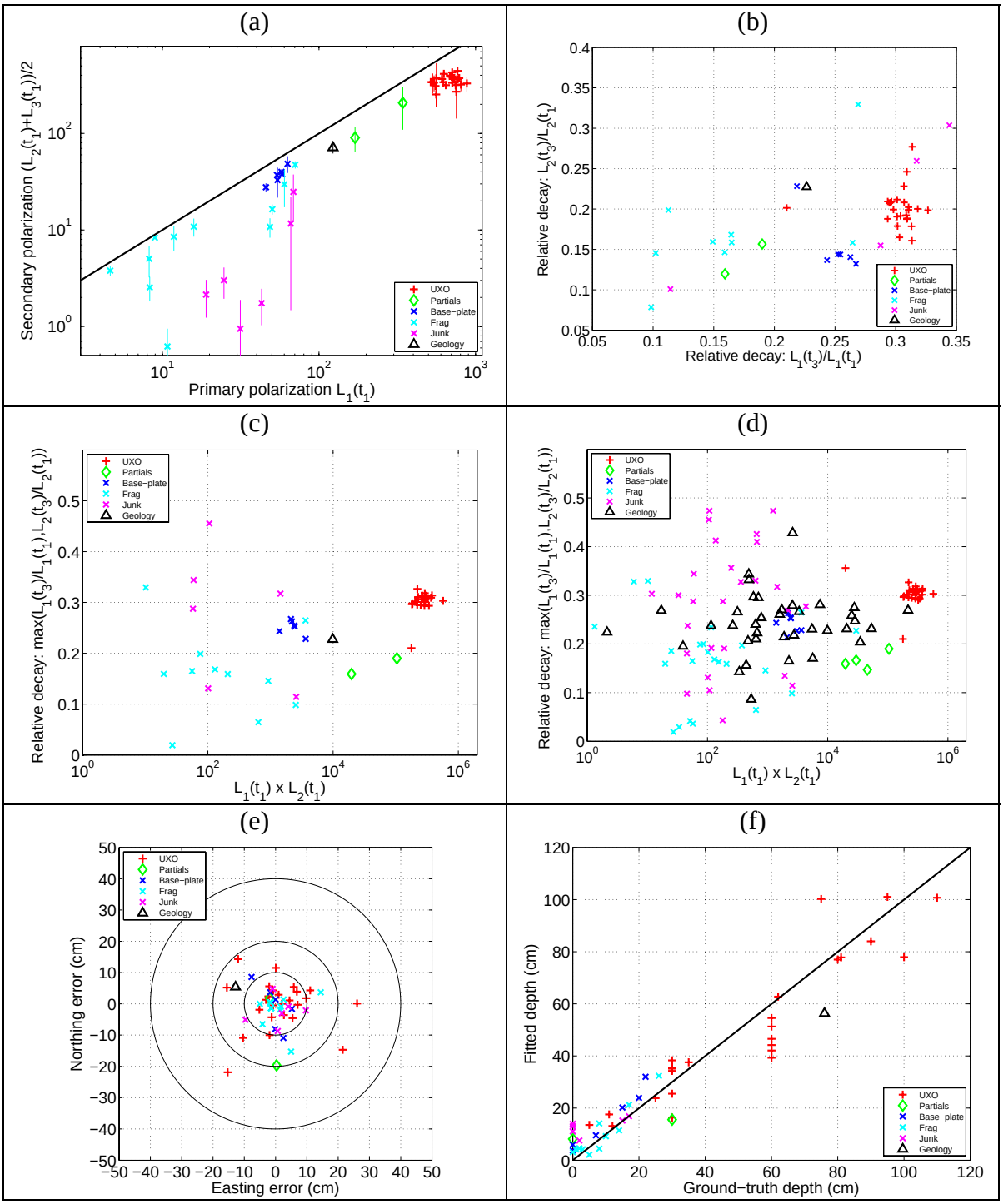



**Figure 21: (a)** Comparison of primary polarization at the first time channel to the secondary polarizations. The extent of the vertical lines are defined by $L_2(t_1)$ and $L_3(t_1)$; **(b)** Comparison of the primary and secondary polarization decay rates (i.e. the ratio between the third and first time channel); **(c)** Comparison of the relative polarization decay with amplitude of the polarizations at the first time channels, **(d)** Repeating the plot in (c) except all fits - including failed fits - are considered; **(e)** Error in estimated dipole locations; **(f)** Predicted versus actual depths.

### b. Analysis of ground-truth data

The rest of the data were inverted using the same procedures as the GPO data. There were 870 inverted anomalies, with 285 having acceptable 3 dipole fit and 591 anomalies with unacceptable fit. The large number of rejected fits is a function of the original picks supplied along with the MTADS EM-61 array data. Many of the picks were geologic in nature (particularly in the SW region) and dipole fits were not attainable. The ground-truth for 136 of the MTADS EM-61 array anomalies was provided by the Program Office in early August, with 51 of the anomalies having valid dipole fits. Twenty-five of the ground-truth items fit were from seeded 4.2" mortars (there were two additional deep rounds in the ground-truth that were not detected, 57 and 1058 @ 120cm as well as two shallower rounds (30cm) that produced unacceptable fits via the cooperative inversion). The dipole parameters and ground-truth information are summarized in Figure 21.

The cooperatively inverted parameters for the 4.2" mortars are more tightly clustered than they were for the single inversion data (Figure 12). This is due to a significant improvement in the accuracy of the position and depth estimates (compare Figures 21(e) and (f) to 13 (a) and b)).

### d. Discrimination strategy for cooperatively inverted MTADS EM61 data

The discrimination strategy for the cooperatively inverted MTADS EM61 data will be the same as that used for the non-cooperatively inverted data. That is, we trained a PNN classifier on $L_1(t_1)$ and the maximum of $L_1(t_1)/ L_3(t_1)$ and $L_2(t_3)/ L_2(t_1)$. Prior to training, the feature vectors were standardized to have 0 mean and unit standard deviation (Figure 21). UXO anomalies with high FOM are well separated from the rest of the items in the ground-truth. One UXO with low FOM lies close to the high FOM boundary between the two classes and would be excavated using the less aggressive cut-off value for low FOM anomalies. Thresholds for construction of the dig-sheet will be the same as those used for the MTADS EM61 data inverted non-cooperatively.

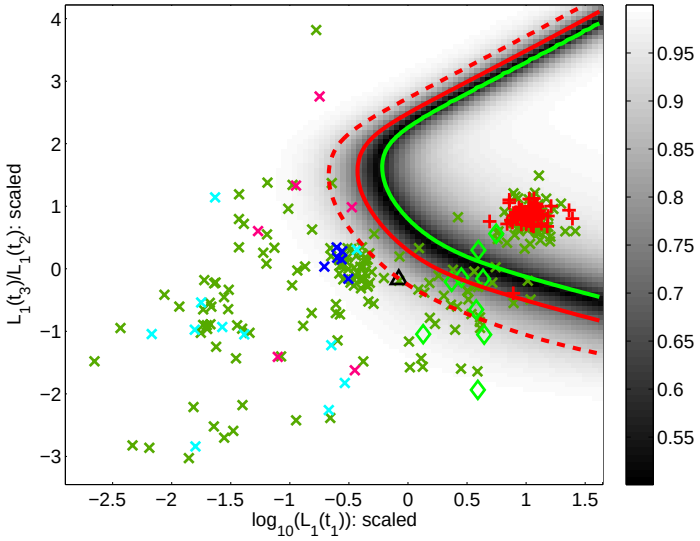



**Figure 21:** PNN classifier trained on feature vectors recovered from MTADS EM61 cooperatively inverted data. Green boundary delineates the high confidence UXO class, the solid red line is the high FOM non-UXO boundary and the dashed red line is the low FOM boundary.

# 6. EM-63 COOPERATIVE DATA

## a. Analysis of Geophysical Prove Out results

The EM-63 data were cooperatively inverted by using dipole location estimates from MTADS magnetics data as *a priori* information. Similar to the cooperatively inverted MTADS EM-61 data, upper and lower constraints on the depth were defined to be twice the estimated variance from the magnetometer data. We opted to use a broader and fixed location constraint of 30 cm in case there were systematic differences between the GPS derived magnetometer and RTS derived EM-63 positions. The noise and mask definitions are the same as for the non-cooperatively inverted data.

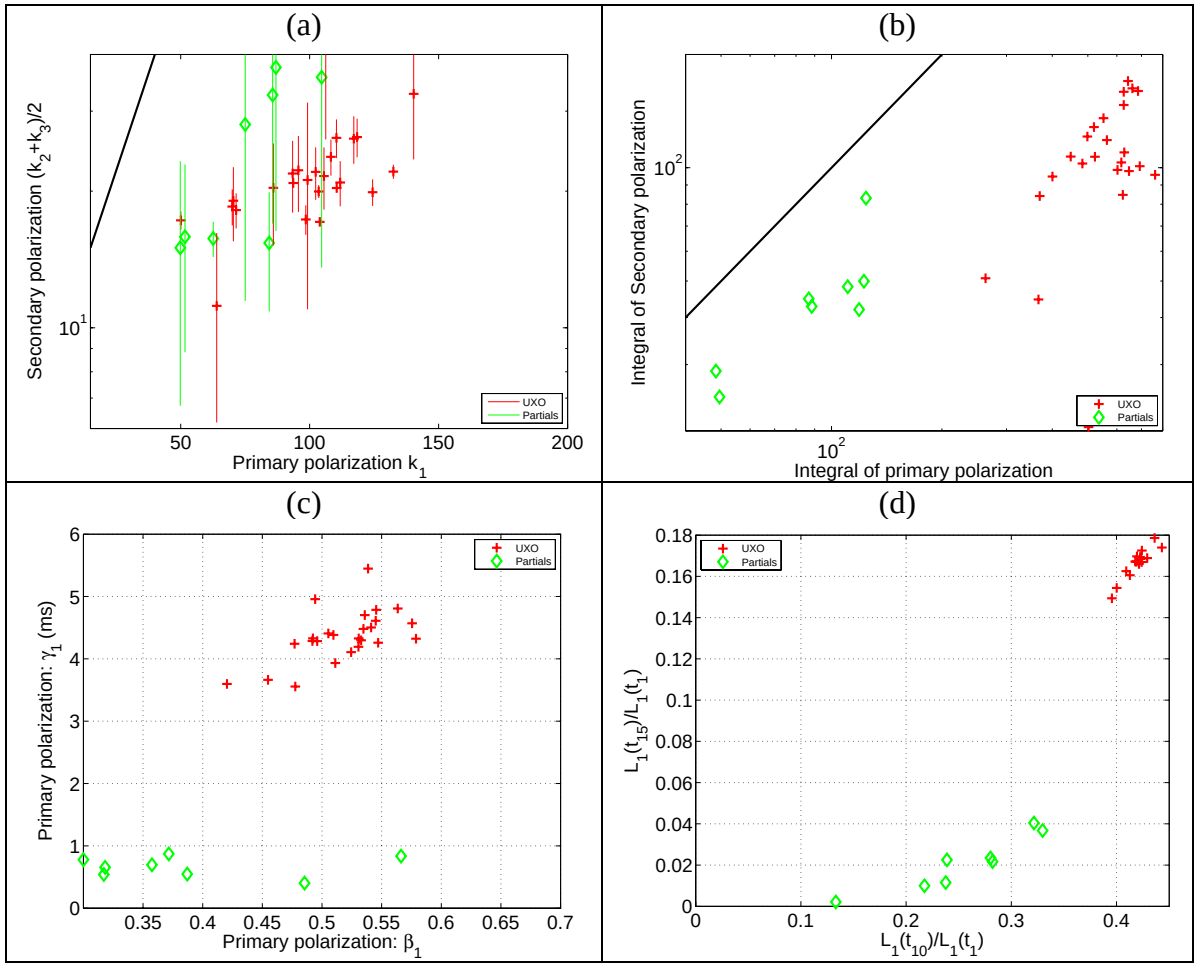



**Figure 22:** Geonics EM-63 cooperative inversion parameter plots for the GPO data. Panels **(a)** and **(b)** compare size-based parameters, and panels **(c)** and **(d)** plot decay parameters.

The 3 polarization model fits were found to be acceptable for all test-plot items with the exception of ten targets (including all 5 targets which were buried at depths greater than 1 m and 5 other target with poor cooperative inversion fits). Parameter plots are shown in Figure 22 and recovered polarization decays in Figure 23. Some parameter values appear to be improved slightly compared to the non-cooperatively inverted data (e.g. integral of polarizations) whereas for others the reverse may be true. For instance, the difference between the secondary

polarizations appears to be larger for UXO's inverted cooperatively. One reason why this might be the case is that the magnetic constraints restrict the amplitude of the polarization tensor and to fit the data, the model is forced to diverge from a 2-dipole model. We do not show position and depth error plots as the results are very similar to the non-cooperatively inverted EM-63 data. Thus, the GPO data indicate that the EM-63 data alone are typically sufficient to constrain the depth.

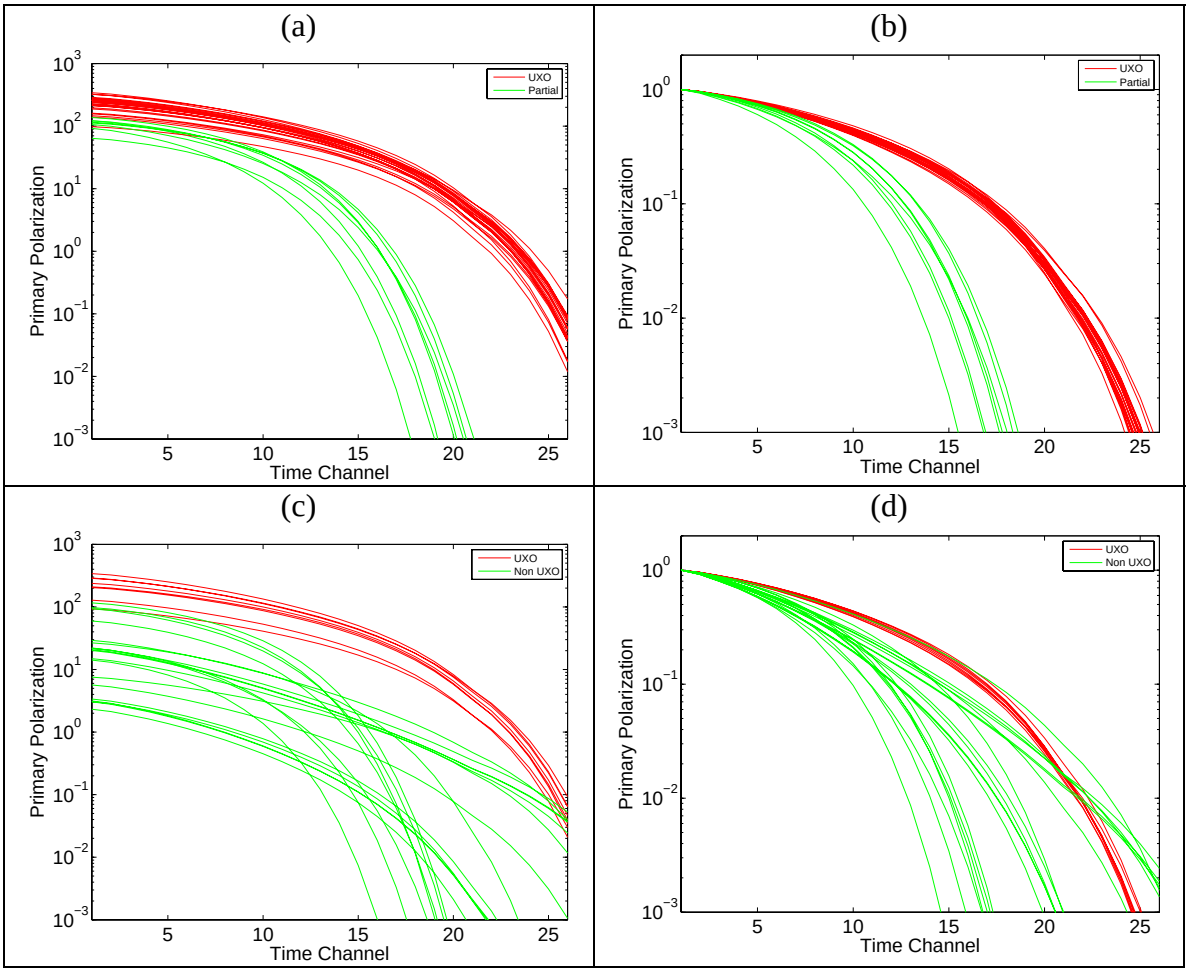



**Figure 23:** Recovered primary polarizations from EM63 cooperative inversions. Panels **(a)** and **(b)** contain raw and normalized recovered primary polarizations from the GPO data, and **(c)** and **(d)** contain raw and normalized recovered primary polarizations from the groundtruth data.

### b. Analysis of ground-truth data

The rest of the data were inverted using the same procedures as the GPO. There were 179 inverted anomalies, with 163 having acceptable 3 dipole fit and 16 having an unacceptable fit. The ground-truth for 28 of the EM-63 anomalies was provided by the Program Office in early August, with 27 of the anomalies having valid dipole fits. Eight of the ground-truth items were from seeded 4.2" mortars. The dipole parameters and ground-truth information are summarized in Figure 25. As with the GPO data there does not appear to be a significant improvement in the clustering of the parameters of the different classes.

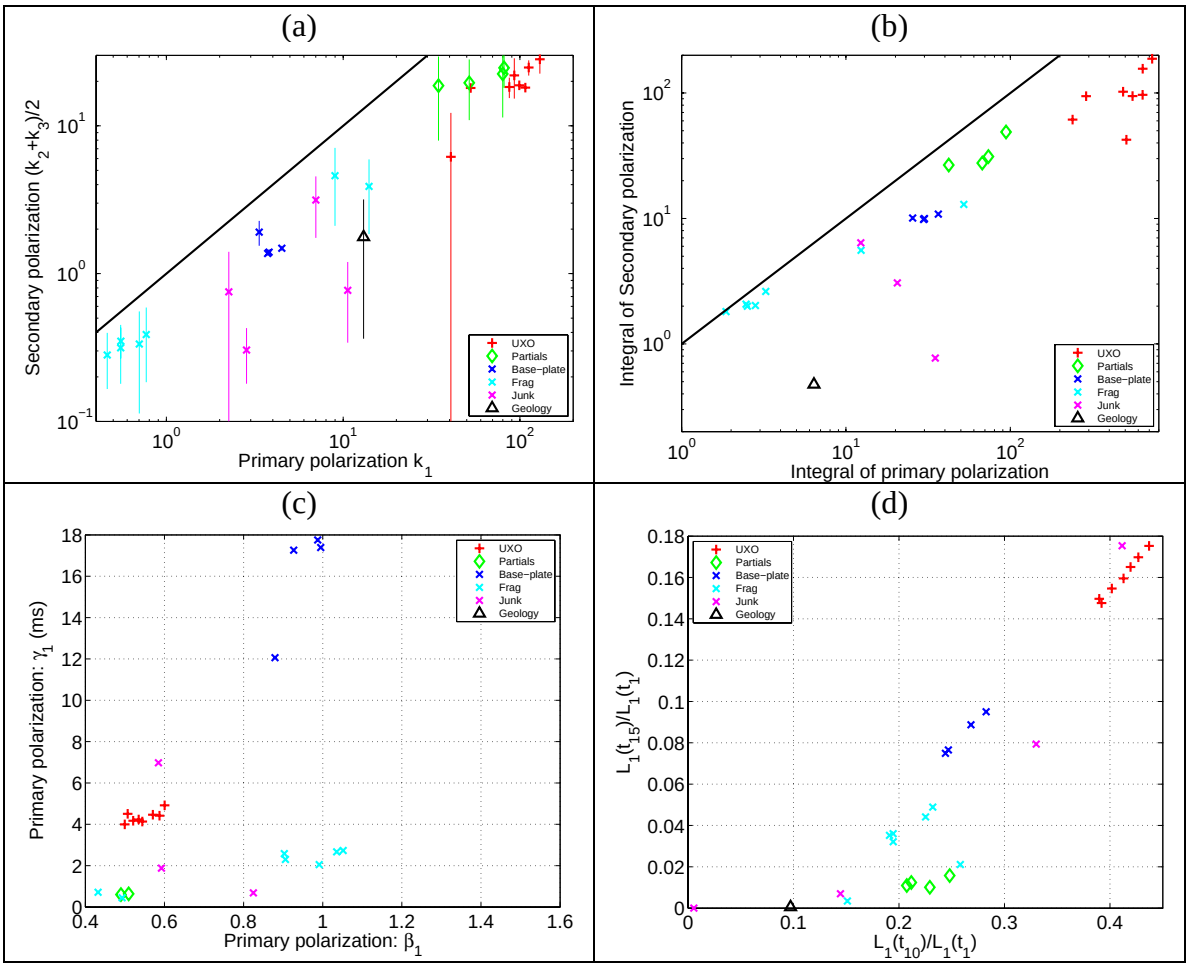



**Figure 25:** Geonics EM-63 inversion parameter plots for the non-GPO data with ground truth. Panels **(a)** and **(b)** compare size-based parameters, and panels **(c)** and **(d)** plot decay characteristics.

### d. Discrimination

A PNN classifier was trained on the GPO and ground-truth items using exactly the same parameters as the EM-63 data inverted non-cooperatively (Figure 26). Qualitatively the classification boundary looks very similar to the non-cooperatively inverted data and we do not expect a significant change in discrimination performance (in contrast to the MTADS data). For digsheet creation, we will use exactly the same thresholds on the PNN probabilities as was used for the non-cooperatively inverted data.

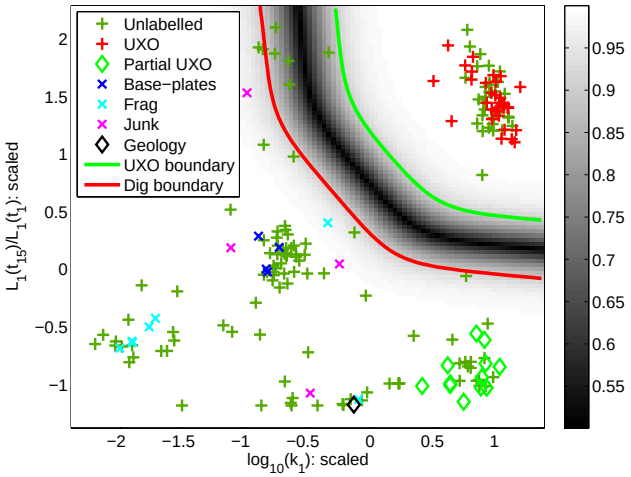



**Figure 26:** PNN classifier applied to $\log_{10}(k_1)$ and $L_1(t_{15})/L_1(t_1)$ for the EM-63 cooperatively inverted data. The color-scale is constructed such that black corresponds to equal probability of membership to either class. The grayscale inside/outside the UXO/non-UXO region indicates increasing probability of membership of the UXO/non-UXO class. Green boundary delineates the region that is high confidence UXO. Red boundary indicates region of non-UXO.